

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 8.04852 Seconds
(without alignments)
418.411 Million cell updates/sec

Title: US-10-792-311-1
Perfect score: 166
Sequence: 1 AKKYAKKEAAKKAYKKAKEAAAEAAKEAAVEA 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	84	50.6	384	B43592	outer membrane prote
2	80.5	48.5	461	T03561	hypothetical prote
3	76	45.8	347	E83525	TolA protein PA097
4	74	44.6	372	G64064	outer membrane int
5	72	43.4	210	A25550	histone H1 - sea u
6	71.5	43.1	211	A28100	histone H1-beta, e
7	70.5	42.5	445	T50972	probable zuotin (i
8	70	42.2	206	S09388	histone H1 - sea u
9	70	42.2	277	T34625	probable NLP/p60 f
10	70	42.2	1015	JC6552	DNA topoisomerase
11	70	42.2	1130	T34081	hypothetical prote
12	69.5	41.9	421	JV0057	TolA protein - Esc
13	67	40.4	394	F90725	membrane spanning
14	67	40.4	394	G85576	membrane spanning
15	67	40.4	1021	G75403	DNA topoisomerase
16	66.5	40.1	217	A26721	histone H1-gamma,
17	66.5	40.1	1052	A44937	kinetoplast-associ
18	66	39.8	231	S59589	histone H1 - Chlam
19	66	39.8	318	H96535	hypothetical prote
20	66	39.8	1390	S51364	sperm tail-specifi
21	65.5	39.5	214	G70673	probable hupB - My
22	65.5	39.5	1528	A60338	surface antigen A
23	65	39.2	101	H59099	hypothetical prote
24	64.5	38.9	185	A32137	histone H1-delta -
25	64	38.6	241	JN0748	histone H1-II - Vo
26	64	38.6	335	T21503	hypothetical prote
27	63.5	38.3	621	A57591	Id-associated prot
28	63	38.0	248	1 HSRU1P	histone H1, gonada
29	63	38.0	261	JN0747	histone H1-I - Vol

30 63 38.0 683 2 H96835
31 63 38.0 735 2 T49622
32 63 38.0 1701 2 T09127
33 62.5 37.7 182 2 S61926
34 62.5 37.7 288 2 T06257
35 62.5 37.7 826 2 A01683
36 62 37.3 206 1 HSTR1R
37 62 37.3 244 2 S40436
38 62 37.3 266 2 T10644
39 62 37.3 376 2 AG0592
40 62 37.3 442 2 T39683
41 62 37.3 628 2 T08942
42 62 37.3 827 2 A01311
43 62 37.3 898 2 A82298
44 62 37.3 899 2 B87553
45 61.5 37.0 217 2 S29309

ALIGNMENTS

RESULT 1

B43592

outer membrane protein TmpB - Treponema phagedenis

C:Species: Treponema phagedenis

C>Date: 30-Jan-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004

C:Accession: B43592

R:Velton, D.B.; Limberger, R.J.; Curci, K.; Malinosky-Rummell, F.; Slivenski, L.; Schoultz, J.; Infect. Immun. 59, 3685-3693, 1991

A:Title: Treponema phagedenis encodes and expresses homologs of the Treponema pallidum Tm

A:Reference number: A43592; MUID:91372983; PMID:1894368

A:Accession: B43592

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-384 <YEL>

A:Cross-references: UNIPROT:P29720; GB:M58563; NID:g155066; PIDN:AAA27480.1; PID:g155067

A>Note: the authors translated the codon TTC for residue 316 as Tyr, and CGA for residue

C:Keywords: membrane protein

Query Match 50.6%; Score 84; DB 2; Length 384;
Best Local Similarity 64.9%; Pred. No. 0.092;
Matches 24; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 1 AKKYAKKEAAKKAYKKAKEAAAEAAKEAAVEA 35
Db 218 AREMAAKEAAKDKAAKEAAEAAKAAEEAAARKAAEEA 254

RESULT 2

T03561

hypothetical protein - Rhodobacter capsulatus

C:Species: Rhodobacter capsulatus

C>Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004

C:Accession: T03561

R:Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M. Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997

A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003.

A:Reference number: Z14955; MUID:97404404; PMID:9256491

A:Accession: T03561

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-461 <VLCS>

A:Cross-references: UNIPROT:O68124; EMBL:AF010496; NID:g3128256; PIDN:AAC16214.1; PID:g3128256

C:Genetics:

A:Map position: 1

Query Match 48.5%; Score 80.5; DB 2; Length 461;
Best Local Similarity 60.0%; Pred. No. 0.25;
Matches 21; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 AKKYAKKEAAKKAYKKAKEAAAEAAKEAAVEA 35
Db 30 AKEVVEKAAAKDAKAAKAAKAR-EEAAAKDAKAAK 63


```
Query Match      42.5%; Score 70.5; DB 2; Length 445;
Best Local Similarity 57.1%; Pred. No. 2.6;
Matches 20; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 2 KKYAKKEKAAYKKEAKAKAA-EAAAKAEAYEA 35
   |||::|||::|||::|||::|||::|||::|||
Db 304 KKRLEKEAAEKKAEEBAKAAEAAKAAEAAKEBEKA 338

RESULT 8
S09388
histone H1 - sea urchin (Parechinus angulosus)
C:Species: Parechinus angulosus (angulate urchin)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S09388
R:Hilli, C.S.; Martin, S.R.; Thomas, J.O.
EMBO J. 8, 2591-2599, 1989
A>Title: A stable alpha-helical element in the carboxy-terminal domain of free and chromatin-associated histone H1
A:Reference number: S09388; MUID:9060019; PMID:2583125
A:Accession: S09388
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-206 <HIL>
A:Cross-references: UNIPROT:Q7M409
C:Superfamily: histone H1
C:Keywords: chromosomal protein

Query Match      42.2%; Score 70; DB 2; Length 206;
Best Local Similarity 56.7%; Pred. No. 1.6;
Matches 17; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 3 KYAKKEKAAYKKEAKAKAAEAAAKAEAA 32
   |||::|||::|||::|||::|||::|||
Db 114 KKAKTSAAKAKKAKARARAAKAKAKAAA 143

RESULT 9
T34625
probable NLP/P60 family secreted protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34625
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21549
A:Accession: T34625
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-277 <SAU>
A:Cross-references: UNIPROT:Q9XAQ3; EMBL:AL078618; PIDN:CAB44532.1; GSPDB:GN00070; SCOEID:
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEID:SC10A7.22

Query Match      42.2%; Score 70; DB 2; Length 277;
Best Local Similarity 54.3%; Pred. No. 2;
Matches 19; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

QY 1 AKKYAKKEKAAYKKEAKAKAAEAAAKAEAYEA 35
   |||::|||::|||::|||::|||::|||
Db 93 AAKQAKTDLAEEKKAAEAAKKAEEAAKAEAAERA 127

RESULT 10
JG6552
DNA topoisomerase (EC 5.99.1.2) - slime mold (Physarum polycephalum)
N:Alternate names: emega-protein; micking-closing enzyme; type I DNA topoisomerase
C:Species: Physarum polycephalum
C>Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: JG6552
R:Czerwinski, R.M.; Lipniacki, A.; Staron, K.
Gene 209, 39-44, 1998
```

```
A>Title: cDNA cloning of Physarum polycephalum DNA topoisomerase I and expression analysis
A:Reference number: JG6552; MUID:98245940; PMID:9583949
A:Accession: JG6552
A:Molecule type: mRNA
A:Residues: 1-1015 <CZE>
A:Cross-references: UNIPROT:Q94705; GB:AF023910; GB:U63217; NID:g2642492; PIDN:AAC14193.1;
C:Genetics:
A:Gene: top1
C:Superfamily: eukaryotic type I DNA topoisomerase
C:Keywords: isomerase
F:974/Active site: Tyr #status predicted

Query Match      42.2%; Score 70; DB 2; Length 1015;
Best Local Similarity 42.9%; Pred. No. 5.9;
Matches 18; Conservative 6; Mismatches 8; Indels 10; Gaps 1;

QY 2 KKYAKKEKAAYKKEK-----AKAKAAEAAAKAEAY 33
   |||::|||::|||::|||::|||::|||::|||
Db 293 KKEVKEDTAKDVKKEVKETPKTTPAKRKAASSSESDF 334

RESULT 11
T34081
hypothetical protein C02F12.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34081
R:Miller, N.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C02F12.
A:Reference number: Z21473
A:Accession: T34081
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1130 <MIL>
A:Cross-references: UNIPROT:Q11102; EMBL:U41545; PIDN:AAA83190.1; CESP:C02F12.7
C:Genetics:
A:Gene: CESP:C02F12.7
A:Introns: 11/2; 118/1; 156/2; 239/2; 308/2; 398/1; 458/3; 591/1; 640/3; 756/2; 790/2; 815/2;

Query Match      42.2%; Score 70; DB 2; Length 1130;
Best Local Similarity 51.4%; Pred. No. 6.4;
Matches 18; Conservative 6; Mismatches 5; Indels 6; Gaps 1;

QY 6 KKEAKAKAYKKEAKAKA-----AEEAAKAEAYE 34
   |||::|||::|||::|||::|||::|||
Db 1016 KKEEASEKRTKAEKAEKAEKAEKAEKAEKAE 1050

RESULT 12
JUV0057
tolA protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: JUV0057; B64810
R:Levengood, S.K.; Webster, R.E.
J. Bacteriol. 171, 6600-6609, 1989
A>Title: Nucleotide sequences of the tolA and tolB genes and localization of their products
A:Reference number: JUV0057; MUID:90078104; PMID:2687247
A:Accession: JUV0057
A:Molecule type: DNA
A:Residues: 1-421 <LEV>
A:Cross-references: UNIPROT:P19934; GB:M28232; NID:g148018; PIDN:AAA24683.1; PID:g148019
A:Experimental source: strain JMW105
A>Note: the authors translated the initiation codon GTG for residue 1 as Val
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64810
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```

A;Residues: 1-421 <BLAT>
A;Cross-references: GB:AE000177; GB:U00096; NID:g1786955; PIDN:AACT3833.1; PID:g1786960;
A;Experimental source: strain K-12, substrain MG1655
C;Comment: tola and tolaB proteins are necessary for collicins E2, E3, A, and K to reach t
C;Genetics:
A;Gene: tolaA
A;Map position: 17 min
A;Start codon: GTG
C;Keywords: nucleotide binding; P-loop; transmembrane protein
F;14-34/Domain: transmembrane #status predicted <MSS>
F;78-301/Domain: helical #status predicted <HSR>
F;355-362/Region: nucleotide-binding motif A (P-loop)
Query Match 41.9%; Score 69.5; DB 2; Length 421;
Best Local Similarity 52.6%; Pred. No. 3.2;
Matches 20; Conservative 4; Mismatches 9; Indels 5; Gaps 1;
QY 3 KYAKKEKAAKKA-----YKKEAKAAEAAAKAEAYEA 35
| : ||| || || ||||| ||| : || : ||
Db 130 KQQAEEAAKAAADAKAKAEADAKAAEAAKKAADA 167
| : ||| || || ||||| ||| : || : ||
RESULT 13
F90725
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90725
R;Hayaashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <HAY>
A;Cross-references: UNIPROT:Q8X965; GB:BA000007; PIDN:BAB34197.1; PID:g13360233; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs0774
Query Match 40.4%; Score 67; DB 2; Length 394;
Best Local Similarity 48.9%; Pred. No. 5.5;
Matches 22; Conservative 5; Mismatches 8; Indels 10; Gaps 2;
QY 1 AKKYAKKE--KAACKAYKK-----EAKAKAAEAAAKAEAYEA 35
||| | : ||| : || : ||| ||||| || : ||
Db 167 AKKKAEEAAKAAAEAKKAAEAAALKKKAAEAAEAAEARKKA 211
| : ||| || || ||||| ||| : || : ||
RESULT 14
G85576
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85576
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamcusis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <STO>
A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:g12513672; PIDN:AAG55075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: tolaA
Query Match 40.4%; Score 67; DB 2; Length 394;
Best Local Similarity 48.9%; Pred. No. 5.5;

Matches 22; Conservative 5; Mismatches 8; Indels 10; Gaps 2;
QY 1 AKKYAKKE--KAACKAYKK-----EAKAKAAEAAAKAEAYEA 35
||| | : ||| : || : ||| ||||| || : ||
Db 167 AKKKAEEAAKAAAEAKKAAEAAALKKKAAEAAEAAEARKKA 211
| : ||| || || ||||| ||| : || : ||
RESULT 15
G75403
DNA topoisomerase I - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: G75403
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; P
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036996; PMID:10567266
A;Accession: G75403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1021 <WHI>
A;Cross-references: UNIPROT:Q9RUL0; GB:AE001993; GB:AE000513; NID:g6459123; PIDN:AAF1094;
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1374
A;Map position: 1
C;Superfamily: bacterial type I DNA topoisomerase
Query Match 40.4%; Score 67; DB 2; Length 1021;
Best Local Similarity 51.4%; Pred. No. 12;
Matches 18; Conservative 3; Mismatches 14; Indels 0; Gaps 0;
QY 1 AKKYAKKEKAACKAYKKAEAKAKAAEAAAKAEAYEA 35
||| | : ||| : || : ||| ||||| || : ||
Db 922 ASKTAPKKAAPKAAKPAKKAAPKKAASKSAKTPAAKA 956
| : ||| || || ||||| ||| : || : ||
Search completed: December 14, 2004, 05:52:11
Job time : 10.0485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 39.2827 Seconds
(without alignments)
512.646 Million cell updates/sec

Title: US-10-792-311-1
Perfect score: 166
Sequence: 1 AKKYAKKEKAAYKKEAKAKAAEAAAYEA 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	50.6	384	1	TMPB TREPH
2	82	49.4	157	1	RS16 CORDI
3	82	49.4	157	2	CAE50058
4	81	48.8	372	2	Q9WWX1
5	81	48.8	372	2	Q88NI6
6	80.5	48.5	461	2	Q68124
7	80	48.2	117	2	Q89DA0
8	79.5	47.9	98	2	Q8WQ44
9	79.5	47.9	111	2	Q8T9R3
10	78.5	47.3	105	2	Q6N503
11	78.5	47.3	105	2	CAE28621
12	78	47.0	395	2	Q6D7F3
13	78	47.0	441	2	Q6N8X8
14	78	47.0	441	2	CAE27215
15	77	46.4	140	1	RS16 SYNXP
16	77	46.4	275	1	MST3 DROHY
17	77	46.4	395	2	Q937K4
18	76	45.8	347	1	TOLA PSEAE
19	75	45.2	244	2	Q9AJX2
20	74.5	44.9	822	2	Q7S5A8
21	74.5	44.9	822	2	CAFO5892
22	74	44.6	189	2	Q8MYC2
23	74	44.6	372	1	TOLA HAEIN
24	73	44.0	156	2	Q8P6U0
25	72	43.4	140	2	Q6MNC2
26	72	43.4	140	2	CAE79230
27	72	43.4	153	1	RS16 BIFLO
28	72	43.4	210	1	H1 LYTFI
29	72	43.4	329	2	Q88RD8
30	71.5	43.1	104	2	Q89JH2
31	71.5	43.1	208	1	DBH_MYCSM

32	71.5	43.1	211	1	H1B_STRPU
33	71.5	43.1	297	2	Q7QC14
34	71	42.8	168	1	RS16 COREF
35	71	42.8	356	2	Q87Y39
36	71	42.8	373	2	Q7W477
37	71	42.8	604	2	Q6F298
38	71	42.8	625	2	Q87V67
39	71	42.8	936	2	Q7UMV5
40	70.5	42.5	44	2	Q711Y2
41	70.5	42.5	44	2	AAQ06805
42	70.5	42.5	445	2	Q9P3Q8
43	70	42.2	140	2	Q7YSZ7
44	70	42.2	206	2	Q7M409
45	70	42.2	277	2	Q9XAQ3

ALIGNMENTS

RESULT 1

ID	TMPB	TREPH	STANDARD	PRT	384 AA
AC	P29720				
DT	01-APR-1993	(Rel. 25, Created)			
DT	01-APR-1993	(Rel. 25, Last sequence update)			
DT	01-OCT-2004	(Rel. 45, Last annotation update)			
DE	Treponemal membrane protein B precursor (Antigen tmpB).				
GN	Name=tmpB;				
OS	Treponema phagedenis.				
OC	Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.				
OX	NCBI_TaxID=162;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Kazan 5;				
RX	MEDLINE=91372983; PubMed=1894368;				
RA	Yellon D.B., Limberger R.J., Curci K., Malinosky-Rummell F.,				
RA	Slivinsky L., Schouls L.M., van Embden J.D.A., Charon N.W.;				
RT	"Treponema phagedenis encodes and expresses homologs of the Treponema				
RT	pallidum tmpA and tmpB proteins.";				
RL	Infect. Immun. 59:3685-3693(1991).				
CC	FUNCTION: Tmp may serve as a porin or transport protein for large				
CC	molecules.				
CC	SUBCELLULAR LOCATION: Outer membrane-associated.				
CC	SIMILARITY: To T.pallidum tmpB.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL; M58563; AA27480.1; -.				
DR	PIR; B43592; B43592.				
DR	InterPro; IPR008941; TPR-like.				
KW	Antigen; Outer membrane; Repeat; Signal.				
FT	SIGNAL 1 21 Potential.				
FT	CHAIN 22 384 Treponemal membrane protein B.				
FT	DOMAIN 151 235 17 X 5 AA tandem repeats of K-A-A-[AKR] -				
FT	[ED].				
FT	REPEAT 151 155 1-1.				
FT	REPEAT 156 160 1-2.				
FT	REPEAT 161 165 1-3.				
FT	REPEAT 166 170 1-4.				
FT	REPEAT 171 175 1-5.				
FT	REPEAT 176 180 1-6.				
FT	REPEAT 181 185 1-7.				
FT	REPEAT 186 190 1-8.				
FT	REPEAT 191 195 1-9.				
FT	REPEAT 196 200 1-10.				
FT	REPEAT 201 205 1-11.				
FT	REPEAT 206 210 1-12.				

```
FT REPEAT      211      215      1-13.
FT REPEAT      216      220      1-14.
FT REPEAT      221      225      1-15.
FT REPEAT      226      230      1-16.
FT REPEAT      231      235      1-17.
FT DOMAIN      236      288      6 X 8 AA tandem repeats of [EA]-A-A-R-X-
FT REPEAT      236      243      2-1.
FT REPEAT      245      252      2-2.
FT REPEAT      254      261      2-3.
FT REPEAT      263      270      2-4.
FT REPEAT      272      279      2-5.
FT REPEAT      281      288      2-6.
SQ SEQUENCE    384 AA; 42677 MW; 6E94CBC74294DB8C CRC64;

Query Match      50.6%; Score 84; DB 1; Length 384;
Best Local Similarity 64.9%; Pred. No. 0.47;
Matches 24; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 1 AKKYAKKEKAAK-KAYKKEAKAKAA-EAAAKEAAEAA 35
   ||||| ||||| ||||| ||||| ||||| |||||
Db 218 AREAAKEKAAKKAYKKEAKAKAAEAAEAAKEAA 254
   ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
RS16 CORDI
ID _RS16 CORDI STANDARD; PRT; 157 AA.
AC P62228; 2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE 30S ribosomal protein S16.
GN Name=rpS6; OrderedLocusNames=DIP1532;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910;
RA Cardeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zeyza A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248358; CAE50058.1; -.
KW Ribosomal protein.
SQ SEQUENCE 157 AA; 16985 MW; BB3FC04EEAD9708A CRC64;

Query Match      49.4%; Score 82; DB 2; Length 157;
Best Local Similarity 64.5%; Pred. No. 0.34;
Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 KKYAKKEKAAKKAYKKEAKAKAAEAAAKEAA 32
   ||||| ||||| ||||| ||||| ||||| |||||
Db 126 KKKAKEEAAKAAEAAEAAAKEEAPAEAA 156
   ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
Q9WWX1
ID Q9WWX1 PRELIMINARY; PRT; 372 AA.
AC Q9WWX1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TolA protein.
GN Name=tolA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96198174; PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope.";
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida.";
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
```

```
RESULT 3
CAE50058
ID CAE50058 PRELIMINARY; PRT; 157 AA.
AC CAE50058;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE 30S ribosomal protein S16.
GN RPS6 OR DIP1532.
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910;
RA Cardeno-Tarraga A.-M., Efstratiou A., Dover L.G., Holden M.T.G.,
RA Pallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zeyza A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RL Nucleic Acids Res. 31:6516-6523(2003).
DR EMBL; BX248358; CAE50058.1; -.
KW Ribosomal protein.
SQ SEQUENCE 157 AA; 16985 MW; BB3FC04EEAD9708A CRC64;

Query Match      49.4%; Score 82; DB 2; Length 157;
Best Local Similarity 64.5%; Pred. No. 0.34;
Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 KKYAKKEKAAKKAYKKEAKAKAAEAAAKEAA 32
   ||||| ||||| ||||| ||||| ||||| |||||
Db 126 KKKAKEEAAKAAEAAEAAAKEEAPAEAA 156
   ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
Q9WWX1
ID Q9WWX1 PRELIMINARY; PRT; 372 AA.
AC Q9WWX1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TolA protein.
GN Name=tolA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96198174; PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope.";
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida.";
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
```

```

RA Ramos-Gonzalez I.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RL EMBL; X74218; CAB50780.1; -.
DR HSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TonB_C.
DR Pfam; PF06519; TOLA; 1.
DR PRINTS; PRO0624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 48.8%; Score 81; DB 2; Length 372;
Best Local Similarity 56.1%; Pred. No. 0.94;
Matches 23; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 1 AKYA-----KKEAKAKYKKEAKAKAAEAAKKAEDAKKAAEEA 35
DB 160 AKKADEAKKAEAEAKKAAEAAEAKKAAEDAKKAAEEA 200

RESULT 6
O68124 PRELIMINARY; PRT; 461 AA.
AC O68124;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Rhodobacter.
OX NCBI_TaxID=1061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SB1003;
RX MEDLINE=97404404; PubMed=9256491;
RA Vlcek C., Paces V., Maltsev N., Paces J., Haselkorn R., Fonstein M.;
RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter
capsulatus SB1003.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).
DR EMBL; AF010496; AAC16214.1; -.
DR PIR; T03561; T03561.
KW Hypothetical protein.
SQ SEQUENCE 461 AA; 49516 MW; DD8DA03418BC0368 CRC64;

Query Match 48.5%; Score 80.5; DB 2; Length 461;
Best Local Similarity 60.0%; Pred. No. 1.3;
Matches 21; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 1 AKYAKKEKAKYKKEAKAKAAEAAKKAEEA 35
DB 30 AKEVVEKAAKDAAKAAKAAKAAKAAKAA 63

RESULT 7
O89DA0 PRELIMINARY; PRT; 117 AA.
AC O89DA0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bll7539 protein.
GN OrderedLocusNames=bll7539;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasanoto S., Watanabe A., Iidesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;

RA Ramos-Gonzalez I.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RL EMBL; X74218; CAB50780.1; -.
DR HSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TonB_C.
DR Pfam; PF06519; TOLA; 1.
DR PRINTS; PRO0624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 48.8%; Score 81; DB 2; Length 372;
Best Local Similarity 56.1%; Pred. No. 0.94;
Matches 23; Conservative 3; Mismatches 9; Indels 6; Gaps 1;

QY 1 AKYA-----KKEAKAKYKKEAKAKAAEAAKKAEEA 35
DB 160 AKKADEAKKAEAEAKKAAEAAEAKKAAEDAKKAAEEA 200

RESULT 5
O88NI6 PRELIMINARY; PRT; 372 AA.
AC O88NI6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biopolymer transport protein TOLA.
GN Name=tOLA; OrderedLocusNames=PPI221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AS016778; AAN66845.1; -.
DR HSP; P50600; 1LR0.
DR TIGR; PPI221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.

```



```

RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CAE28621.1; -.
KW Hypothetical protein.
SQ
SEQUENCE 105 AA; 11042 MW; CEDB59B3D937B980 CRC64;

Query Match 47.3%; Score 78.5; DB 2; Length 105;
Best Local Similarity 63.2%; Pred. No. 0.56;
Matches 24; Conservative 4; Mismatches 7; Indels 3; Gaps 2;

QY 1 AKKYAKKEKA-AKKAYK--KEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 33 AKKVAKVKRAGAKAATKAGKGAAGAAKKAAGKAAKKA 70

RESULT 12
Q6D7F3 PRELIMINARY; PRT; 395 AA.
ID AC Q6D7F3;
AC Q6D7F3;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE TolA protein.
GN Name=tolA; ORFNames=ECA1372;
OS Erwinia carotovora subsp. atroseptica SCRI1043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCRI1043;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
RA Aleva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.,
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG74282.1; -.
SQ SEQUENCE 395 AA; 41698 MW; 2DB1DAA64E984199 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 395;
Best Local Similarity 57.1%; Pred. No. 2.1;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 AKQVAAAAAKQAEEAEKAEAAKQKAAETA 207

RESULT 13
Q6N8X8 PRELIMINARY; PRT; 441 AA.
ID AC Q6N8X8;
AC Q6N8X8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE OmpA/MotB domain, possible porin precursor.
OS Rhodospseudomonas palustris.
GN Ordered locus Names=RPAL1774;
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,

```

```

RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
CC -!- SIMILARITY: Belongs to the ompA family.
DR EMBL; BX572598; CAE27215.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005741; C:mitochondrial outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR InterPro; IPR006664; Bac OmpA.
DR InterPro; IPR006665; OmpA/MotB.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PRO1021; OMPADOMAIN.
DR ProDom; PD000930; OmpA/MotB; 1.
KW Complete proteome; Porin; Signal.
FT SIGNAL 1 27 Potential.
SQ SEQUENCE 441 AA; 44811 MW; F6BB66A4AFE183A0 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 441;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 AKEVAAKEAAKAAEAEEAAKQAAADEAAKKADEA 293

RESULT 14
CAE27215 PRELIMINARY; PRT; 441 AA.
ID AC CAE27215;
AC CAE27215;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE OmpA/MotB domain, possible porin precursor.
GN RPAL1774.
OS Rhodospseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodospseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodospseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572598; CAE27215.1; -.
KW Porin; Signal.
FT SIGNAL 1 27 Potential.
SQ SEQUENCE 441 AA; 44811 MW; F6BB66A4AFE183A0 CRC64;

Query Match 47.0%; Score 78; DB 2; Length 441;
Best Local Similarity 57.1%; Pred. No. 2.3;
Matches 20; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 AKKYAKKEKA-AKKAYKKEAKAKAAEAEEAAKEAAVEA 35
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 259 AKEVAAKEAAKAAEAEEAAKQAAADEAAKKADEA 293

RESULT 15
RS16 SYNXP STANDARD; PRT; 140 AA.
ID AC RS16 SYNXP;
AC Q7TTU5;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE 30S ribosomal protein S16.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 38.692 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-1
Perfect score: 166
Sequence: 1 AKKYAKKEKAAYKAYKEAKAKAAAEAAKEAAEYA 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	166	100.0	35	3 AAY82571	Aay82571 Copolymer
2	108	65.1	45	3 AAY82572	Aay82572 Copolymer
3	108	65.1	86	3 AAY82576	Aay82576 Copolymer
4	105.5	63.6	66	3 AAY82574	Aay82574 Copolymer
5	105	63.3	109	3 AAY82577	Aay82577 Copolymer
6	104.5	63.0	56	3 AAY82573	Aay82573 Copolymer
7	104.5	63.0	77	3 AAY82575	Aay82575 Copolymer
8	86.5	52.1	427	3 AAG70868	Aag70868 C albican
9	86.5	52.1	427	8 ADP99051	Adp99051 C. albica
10	82	49.4	157	6 ABU26064	Abu26064 Protein e
11	81	48.8	372	6 ABU40185	Abu40185 Protein e
12	76	45.8	154	2 AAR06445	Aar06445 Recombina
13	76	45.8	347	6 ABU38313	Abu38313 Protein e
14	76	45.8	347	6 ABJ18771	Abj18771 Pseudomon
15	76	45.8	407	7 ABO80835	Abc80835 Pseudomon
16	74.5	44.9	80	5 ABG71044	Abg71044 Tumour ne
17	74	44.6	372	5 ABG80418	Abg80418 Haemophil
18	74	44.6	372	7 ABO23507	Abc23507 Haemophil
19	73.5	44.3	40	3 AAB08170	Aab08170 Peptide m
20	72	43.4	80	7 ADE10657	Adel10657 Structura
21	72	43.4	80	7 ADE10656	Adel10656 Structura
22	72	43.4	80	7 ADE10606	Adel10606 Structura
23	72	43.4	80	7 ADE10607	Adel10607 Structura
24	72	43.4	80	8 ADKI5676	Adki5676 Library f
25	72	43.4	80	8 ADKI5675	Adki5675 Library f

26	72	43.4	80	8 ADKI5626	Adki5626 Nucleatin
27	72	43.4	80	8 ADKI5625	Adki5625 Nucleatin
28	72	43.4	85	7 ADE10608	Adel10608 Structura
29	72	43.4	85	7 ADE10658	Adel10658 Structura
30	72	43.4	85	8 ADKI5627	Adki5627 Nucleatin
31	72	43.4	85	8 ADKI5677	Adki5677 Library f
32	72	43.4	153	5 ABP65446	Abp65446 Bifidobac
33	72	43.4	329	6 ABU39775	Abu39775 Protein e
34	71	42.8	31	8 ADN11701	Adn11701 Peptide l
35	71	42.8	106	2 AAR06446	Aar06446 Recombina
36	71	42.8	336	6 ABU42038	Abu42038 Protein e
37	69.5	41.9	145	4 AAM25508	Aam25508 Human pro
38	69.5	41.9	165	4 AAG91997	Aag91997 C glutami
39	69.5	41.9	421	6 ABU28559	Abu28559 Protein e
40	69	41.6	741	4 ABB62028	Abb62028 Drosophi
41	67.5	40.7	32	2 AAR90180	Aar90180 Polycatio
42	67.5	40.7	32	2 AAW06686	Aaw06686 Protamine
43	67.5	40.7	32	3 AAB08168	Aab08168 Peptide m
44	67.5	40.7	33	2 AAR90181	Aar90181 Polycatio
45	67.5	40.7	33	2 AAW06688	Aaw06688 Protamine

ALIGNMENTS

RESULT 1
AAY82571
ID AAY82571 standard; peptide; 35 AA.

AC AAY82571;
XX
DT 28-JUL-2000 (first entry)
XX
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:1.
XX
KW Copolymer; molecular weight marker; TV-marker; immune disease;
KW glatiramer acetate; autoimmune disease; ankiarthritic; neuroprotective;
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
KW pemphigus vulgaris; systemic lupus erythematosus.
XX
OS Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA) YEDA RES & DEV CO LTD.
(TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for
glatiramer acetate and for treatment and prevention of immune diseases.
PT Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
weight TV-marker polypeptides from the present invention. The present
invention describes polypeptides (I) for determining the molecular weight
of a copolymer (CP), which has an identified molecular weight and an
amino acid composition corresponding to the copolymer. The polypeptides


```

XX OS Unidentified.
XX PN WO200018794-A1.
XX XX
XX PD 06-APR-2000.
XX XX
XX PF 24-SEP-1999; 99WO-US022402.
XX XX
XX PR 25-SEP-1998; 98US-0101693P.
XX XX
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (TEVA-) TEVA PHARM USA INC.
XX XX
XX PI Gad A, Lis D;
XX XX
XX DR WPI; 2000-317499/27.
XX XX
XX PT Copolymer 1 related polypeptides used as molecular weight markers for
XX PT glatiramer acetate and for treatment and prevention of immune diseases.
XX PS Claim 10; Page 14; 72pp; English.
XX CC
XX CC AA82571 to AA82577 represent specifically claimed copolymer molecular
XX CC weight TV-marker polypeptides from the present invention. The present
XX CC invention describes polypeptides (I) for determining the molecular weight
XX CC of a copolymer (CP), which has an identified molecular weight and an
XX CC amino acid composition corresponding to the copolymer. The polypeptides
XX CC of the invention are used as molecular weight markers for glatiramer
XX CC acetate related tetrapolymers. The polypeptides may also be used for
XX CC treating and preventing immune diseases in a mammal. Autoimmune diseases
XX CC which may be treated include either cell-mediated or antibody-mediated
XX CC diseases. Such diseases include arthritic conditions, demyelinating
XX CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
XX CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
XX CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
XX CC disease, chronic immune thrombocytopenia purpura, colitis, contact
XX CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
XX CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
XX CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
XX CC mediated diseases which can be treated include host-versus-graft disease,
XX CC graft-versus-host disease, and delayed-type hypersensitivity. The
XX CC polypeptides of the invention have defined molecular weights and physical
XX CC properties which are analogous to glatiramer acetate molecules, which
XX CC makes them ideal for use as molecular weight markers
XX SQ Sequence 77 AA;

Query Match 63.0%; Score 104.5; DB 3; Length 77;
Best Local Similarity 65.8%; Pred. No. 6.4e-05;
Matches 25; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 1 AKKYAKKEKAAYKAY---KKEAKKAATAAAKAAKAAAYEA 35
DB ||||| ||| ||| : : |||:|||||||
40 AKKYAKAAKAEKKEAYAAAEAKYKAAKAAKAAKAAAYEA 77

RESULT 8
AAG70868
ID AAG70868 standard; protein; 427 AA.
XX
XX AC AAG70868;
XX XX
XX DT 27-JUL-2001 (first entry)
XX XX
XX DE C albicans apoptosis associated protein #48.
XX XX
XX KW Yeast; fungus; apoptosis; infection; proliferative disease; vaccine;
XX KW autoimmune disease; ischaemia; neurodegeneration.
XX XX
XX OS Candida albicans.
XX PN WO200102550-A2.

```

```

XX PD 11-JAN-2001.
XX XX
XX PF 03-JUL-2000; 2000WO-BE000077.
XX XX
XX PR 01-JUL-1999; 99EP-00870141.
XX XX
XX PA (JANC ) JANSSEN PHARM NV.
XX XX
XX PI Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
XX PI Nelissen BJM, Reekmans RJ;
XX XX
XX DR WPI; 2001-367042/38.
XX DR N-PSDB; AAH29904.
XX XX
XX PT Yeast and fungal nucleic acids encoding proteins involved in a pathway
XX PT leading to programmed cell death, useful for treating proliferative
XX PT disorders, yeast and fungal infections, or for preventing apoptosis in
XX PT certain diseases.
XX PS Claim 24; Fig 2; 218pp; English.
XX CC
XX CC The present invention provides the protein and coding sequences of a
XX CC number of apoptosis associated proteins from the yeast Saccharomyces
XX CC cerevisiae and the fungus Candida albicans. These can be used to identify
XX CC treatments for fungal and yeast infections, for proliferative diseases
XX CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
XX CC and neurodegeneration. The present sequence is one of the C. albicans
XX CC proteins of the invention
XX SQ Sequence 427 AA;

Query Match 52.1%; Score 86.5; DB 4; Length 427;
Best Local Similarity 65.7%; Pred. No. 0.044;
Matches 23; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 2 KKYAKKEKAAYKAYKKEAKAKAA-EAAAKEAAYEA 35
DB :||||||| : ||| |||||:|
283 KEKAKKEKAAYKKEKSGSRKAABEAAAKAAEEA 317

RESULT 9
ADP99051
ID ADP99051 standard; protein; 427 AA.
XX
XX AC ADP99051;
XX XX
XX DT 23-SEP-2004 (first entry)
XX XX
XX DE C. albicans specific gene, orf19.2709, protein sequence.
XX XX
XX KW Diploid fungal cell; allele; gene disruption cassette;
XX KW promoter replacement fragment; antifungal; fungicide; gene therapy;
XX KW infection; Candida albicans.
XX OS Candida albicans.
XX PN WO2004056965-A2.
XX XX
XX PD 08-JUL-2004.
XX XX
XX PF 19-DEC-2003; 2003WO-US040618.
XX XX
XX PR 19-DEC-2002; 2002US-0434832P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX PA (ELIT-) ELITRA CANADA LTD.
XX XX
XX PI Roemer T, Jiang B, Boone C, Bussey H;
XX DR WPI; 2004-500296/47.
XX DR N-PSDB; ADP98741.
XX XX

```

Constructing a strain of diploid fungal cells in which both alleles of a gene are modified comprises modifying the alleles of a gene in the fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment.

Claim 44; SEQ ID NO 7226; 163pp; English.

The invention relates to a novel method for constructing a strain of diploid fungal cells in which both alleles of a gene are modified. The method comprises modifying the alleles of a gene in diploid fungal cells by recombination using a gene disruption cassette and a promoter replacement fragment. The invention further comprises: assembling a collection of diploid fungal cells each of which comprises modified alleles of a different gene; a strain of diploid fungal cells comprising modified alleles of a gene, where the first allele of the gene is inactivated by a gene disruption cassette comprising a nucleotide sequence encoding an expressible selectable marker; and the expression of the second allele of the gene is regulated by a heterologous promoter that is operably linked to the coding region of the second allele of the gene, and where the gene encodes the polypeptide mentioned above; a collection of diploid fungal strains comprising the diploid strains cited above, where substantially all the different genes that encode the above amino acid sequences are modified and are present in different diploid strains in the collection; a nucleic acid molecule microarray comprising nucleic acid molecules, where each nucleic acid molecule comprises a nucleotide sequence that is hybridizable to a target nucleotide sequence comprising any of the 310 nucleotide sequences listed in the specification (ADP98516-ADP98825); identifying a gene that is essential to the survival or growth of a fungus, that contributes to the virulence and/or pathogenicity of a fungus, or that contributes to the resistance of a diploid fungus to an antifungal agent; identifying an antifungal agent that inhibits the growth of a diploid fungus, or a therapeutic agent for treatment of a mammalian disease; correlating changes in the levels of proteins or gene transcripts with the inhibition of growth or proliferation of a diploid fungal cell; a purified or isolated nucleic acid molecule comprising a nucleotide sequence encoding a gene product required for proliferation of *Candida albicans*, where the gene product consists of any of the above-mentioned amino acid sequences; a vector comprising a promoter operably linked to the nucleic acid molecule cited above; a host cell containing the vector; a purified or isolated polypeptide comprising any of the 61 amino acid sequences given in the specification (ADP96718-ADP96778); a fusion protein comprising a fragment of a first polypeptide fused to a second polypeptide, the fragment consisting of at least 6 consecutive residues of any of ADP98826-ADP99135; producing a polypeptide; identifying a compound which modulates the activity of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825; eliciting an immune response in an animal; a strain of *Candida albicans*, where a first allele of a gene comprising any of ADP98516-ADP98825 is inactive and a second allele of the gene is under the control of a heterologous promoter; identifying a compound or binding partner that binds to the polypeptide comprising any of ADP98826-ADP99135, or its fragment; identifying a compound having the ability to inhibit growth or proliferation of *Candida albicans*; inhibiting growth or proliferation of *Candida albicans* cells; manufacturing an antimycotic compound; treating an infection of a subject by *Candida albicans*; preventing or containing contamination of an object by *Candida albicans*, or for preventing or inhibiting formation on a surface of a biofilm comprising *Candida albicans*; a pharmaceutical composition comprising a therapeutic amount of an agent which reduces the activity or level of a gene product encoded by a nucleic acid comprising any of ADP98516-ADP98825 in a pharmaceutical carrier; an antibody preparation which binds the polypeptide; methods for evaluating a compound against a target gene product encoded by any of ADP98516-ADP98825; identifying an antimycotic compound; a computer or a computer readable medium that comprises at least one of the nucleotide sequences mentioned in the specification or at least one amino acid sequence selected from ADP98826-ADP99135; a method assisted by a computer for identifying a putatively essential gene of a fungus; and a protein array comprising proteins, where at least one protein comprises an amino acid sequence or a portion of an amino acid sequence selected from ADP98516-ADP98825. The novel methods and compositions have fungicide activity. The compositions may be used in gene therapy. The composition and methods are useful for drug screening purposes or for diagnosing, preventing or treating infections associated

CC with *Candida albicans*. These may also be used for constructing strains useful for identification and validation of gene products as effective targets for therapeutic intervention, for identifying and validating gene products as effective targets for therapeutic intervention, and for collecting identified essential genes. This sequence represents the protein of a *Candida albicans* fungal specific gene of the invention. CC NOTE: This sequence was downloaded from an electronic sequence listing CC provided on the WIPO website.

XX Sequence 427 AA;

Query Match 52.1%; Score 86.5; DB 8; Length 427;

Best Local Similarity 65.7%; Pred. No. 0.044;

Matches 23; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 2 KKYAKKEKAAYKKEAKAKAA-EAAAKEAAYEA 35

DB 283 KEKAKEKAAYKKEGSRKAEEAAAKAAEEA 317

RESULT 10

ABU26064
ID ABU26064 standard; protein; 157 AA.

AC ABU26064;

DT 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #11591.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX *Corynebacterium diphtheriae*.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-029926/02.

DR N-PSDB; ACA29934.

XX New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 53988; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation or screening for homologous nucleic acids
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 157 AA;

Query Match 49.4%; Score 82; DB 6; Length 157;
Best Local Similarity 64.5%; Pred. No. 0.054;
Matches 20; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 2 KKYAKKEAAKAYKKEAKAAEAAAEAA 32
|||:||||:|||||:|||||:|||||
Db 126 KKKAKEAAKAAEAAEAAEAAEAPAEAA 156

RESULT 11
ABU40185
ID ABU40185 standard; protein; 372 AA.
XX
AC ABU40185;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #25712.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Pseudomonas putida.
XX
PN WO20027183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA44055.
XX

PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 68109; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 372 AA;

Query Match 48.8%; Score 81; DB 6; Length 372;
Best Local Similarity 56.1%; Pred. No. 0.17;
Matches 23; Conservative 3; Mismatches 9; Indels 6; Gaps 1;
QY 1 AKKYA-----KCKKAACKAYKKEAKAAEAAAEAAEAA 35
|||:||||:|||||:|||||:|||||:|||||
Db 160 AKKXADEAKKABEEAKKAAAEAAEAKKAAEDAKKAAEEA 200

RESULT 12
AAR06445
ID AAR06445 standard; protein; 154 AA.
XX
AC AAR06445;
XX
DT 25-MAR-2003 (revised)
DT 03-JAN-1991 (first entry)
XX
DE Recombinant copolymer 1-77, myelin basic protein analogue.
XX
KW Recombinant copolymer 1, COP-1-77; myelin basic protein; MBP;
KW immunological activity; autoimmune encephalomyelitis; multiple sclerosis.
XX
OS Synthetic.
XX
PN EP383620-A.
XX
PD 22-AUG-1990.
XX
PF 16-FEB-1990; 90EP-00301700.
XX
PR 17-FEB-1989; 89US-00312541.
PR 07-FEB-1990; 90US-00473845.
XX
PA (REPK) REPLIGEN CORP.
XX
PI Cook KS;
XX
DR WPI; 1990-255848/34.
DR N-PSDB; AAQ05664.
XX

PT Producing genes encoding random polymers of aminoacid(s) - for producing

The invention relates to *Pseudomonas aeruginosa* polypeptides and the polynucleotides encoding them. The sequences are useful in diagnosis and therapy of pathological conditions, as molecular targets for diagnostics and prophylaxis and treatment of pathological conditions resulting from a bacterial infection, for evaluating a compound, such as a polypeptide, for the ability to bind a *P. aeruginosa* nucleic acid, as components of effective antibacterial targets, as targets for antibacterial drugs,

Search completed: December 14, 2004, 06:01:08
Job time : 46.692 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:52:22 ; Search time 86.8354 Seconds
(without alignments)
143.965 Million cell updates/sec

Title: US-10-792-311-1

Perfect score: 166

Sequence: 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	166	100.0	35	9	US-09-816-989A-1	Sequence 1, Appli
2	108	65.1	45	9	US-09-816-989A-2	Sequence 2, Appli
3	108	65.1	86	9	US-09-816-989A-6	Sequence 6, Appli
4	105.5	63.6	66	9	US-09-816-989A-4	Sequence 4, Appli
5	105	63.3	109	9	US-09-816-989A-7	Sequence 7, Appli
6	104.5	63.0	56	9	US-09-816-989A-3	Sequence 3, Appli
7	104.5	63.0	77	9	US-09-816-989A-5	Sequence 5, Appli
8	82	49.4	157	15	US-10-282-122A-53988	Sequence 53988, A
9	81	48.8	372	15	US-10-282-122A-68109	Sequence 68109, A
10	76	45.8	347	14	US-10-127-032-120	Sequence 120, App
11	76	45.8	347	15	US-10-282-122A-66237	Sequence 66237, A
12	74	44.6	372	16	US-09-820-843A-8	Sequence 8, Appli
13	74	44.6	372	16	US-10-467-421-16	Sequence 16, Appli

14	72	43.4	80	14	US-10-177-725-13	Sequence 13, Appl
15	72	43.4	80	14	US-10-177-725-14	Sequence 14, Appl
16	72	43.4	80	14	US-10-177-725-63	Sequence 63, Appl
17	72	43.4	80	14	US-10-177-725-64	Sequence 64, Appl
18	72	43.4	80	14	US-10-393-449-13	Sequence 13, Appl
19	72	43.4	80	14	US-10-393-449-14	Sequence 14, Appl
20	72	43.4	80	14	US-10-393-449-63	Sequence 63, Appl
21	72	43.4	80	14	US-10-393-449-64	Sequence 64, Appl
22	72	43.4	80	14	US-10-177-725-15	Sequence 15, Appl
23	72	43.4	85	14	US-10-177-725-65	Sequence 65, Appl
24	72	43.4	85	14	US-10-393-449-15	Sequence 15, Appl
25	72	43.4	85	14	US-10-393-449-65	Sequence 65, Appl
26	72	43.4	329	15	US-10-282-122A-67699	Sequence 67699, A
27	71	42.8	336	15	US-10-282-122A-69962	Sequence 69962, A
28	70.5	42.5	189	16	US-10-767-701-60774	Sequence 60774, A
29	70	42.2	1130	14	US-10-369-493-6751	Sequence 6751, Ap
30	69.5	41.9	146	15	US-10-296-115-1023	Sequence 1023, Ap
31	69.5	41.9	165	9	US-09-738-626-5751	Sequence 5751, Ap
32	69.5	41.9	421	15	US-10-282-122A-56483	Sequence 56483, A
33	67	40.4	258	14	US-10-156-761-9957	Sequence 9957, Ap
34	66.5	40.1	79	14	US-10-177-725-20	Sequence 20, Appl
35	66.5	40.1	79	14	US-10-393-449-20	Sequence 20, Appl
36	66.5	40.1	582	9	US-09-919-497-100	Sequence 100, App
37	66	39.8	120	16	US-10-767-701-45061	Sequence 45061, A
38	66	39.8	269	17	US-10-425-115-27374	Sequence 27374, A
39	66	39.8	526	15	US-10-282-122A-53742	Sequence 53742, A
40	66	39.8	685	14	US-10-369-493-3684	Sequence 3684, Ap
41	66	39.8	1593	15	US-10-282-122A-65262	Sequence 65262, A
42	65.5	39.5	79	14	US-10-177-725-16	Sequence 16, Appl
43	65.5	39.5	79	14	US-10-393-449-16	Sequence 16, Appl
44	65.5	39.5	212	15	US-10-282-122A-61735	Sequence 61735, A
45	65.5	39.5	214	14	US-10-229-567-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-09-816-989A-1
; Sequence 1, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKI
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 35
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-1

Query Match 100.0%; Score 166; DB 9; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.8e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

Db 1 AKKYAKKEKAAYKKEAKAAEAAAEAAAYEA 35

RESULT 2

```
US-09-816-989A-2
; Sequence 2, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-2
Query Match 65.1%; Score 108; DB 9; Length 45;
Best Local Similarity 64.4%; Pred. No. 5.2e-05;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;
QY 1 AKKYAKKEKA-AKKAYK-----KEAKAAEAEEAAEAYEA 35
Db 1 AKKYAKKAERKAYKAAEAKKAAKAYEKAEEAAEAAEAYEA 45

RESULT 3
US-09-816-989A-6
; Sequence 6, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-6
Query Match 65.1%; Score 108; DB 9; Length 86;
Best Local Similarity 57.4%; Pred. No. 0.0001;
Matches 27; Conservative 2; Mismatches 6; Indels 12; Gaps 1;
QY 1 AKKYAKKEKA-AKKAYK-----KKEAKAAEAEEAAEAYEA 35
Db 40 AKKYAKKAERKAYEAAAEAKYKAAEAKKAYKAAEAAEAAEAYEA 86

RESULT 4
US-09-816-989A-4
; Sequence 4, Application US/09816989A
; Patent No. US20020115103A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-4
Query Match 63.6%; Score 105.5; DB 9; Length 66;
Best Local Similarity 43.9%; Pred. No. 0.00014;
Matches 29; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
QY 1 AKKYAKKEKA-AKKAYKKEA-----KAKAAEAEEAAK 29
Db 1 AKKYAKKEKAYAKKAAEAKKAAKAAKAAKAAEAKKAAEAKYKAAEAAEAAK 60
QY 30 EAAEYA 35
Db 61 EAAEYA 66

RESULT 5
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7
Query Match 63.3%; Score 105; DB 9; Length 109;
Best Local Similarity 62.2%; Pred. No. 0.00027;
Matches 28; Conservative 2; Mismatches 5; Indels 10; Gaps 2;
QY 1 AKKYAKKEKA-AKKAYK-----KKEA-----KAKAAEAEEAAEAYEA 35
Db 65 AKKYAKKAERKAYEAAAEAKKAAEAKKAAEAKKAAEAKKAAEAAEAYEA 109

RESULT 6
US-09-816-989A-3
; Sequence 3, Application US/09816989A
```


This Page Blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	166	100.0	35	4	US-09-405-743A-1	Sequence 1, Appli
2	108	65.1	45	4	US-09-405-743A-2	Sequence 2, Appli
3	108	65.1	46	4	US-09-405-743A-6	Sequence 6, Appli
4	105.5	63.6	66	4	US-09-405-743A-4	Sequence 4, Appli
5	105	63.3	109	4	US-09-405-743A-7	Sequence 7, Appli
6	104.5	63.0	56	4	US-09-405-743A-3	Sequence 3, Appli
7	104.5	63.0	77	4	US-09-405-743A-5	Sequence 5, Appli
8	86.5	52.1	176	4	US-09-248-796A-18922	Sequence 18922, A
9	76	45.8	407	4	US-09-252-991A-29581	Sequence 29581, A
10	67.5	40.7	32	1	US-08-152-488-13	Sequence 13, Appl
11	67.5	40.7	32	1	US-08-303-025-15	Sequence 15, Appl
12	67.5	40.7	32	1	US-08-677-304-13	Sequence 13, Appl
13	67.5	40.7	32	2	US-08-436-703B-2	Sequence 2, Appli
14	67.5	40.7	33	3	US-08-303-025-16	Sequence 16, Appl
15	67.5	40.7	33	2	US-08-436-703B-4	Sequence 4, Appli
16	66.5	39.5	116	4	US-09-019-497-100	Sequence 100, App
17	65.5	39.5	214	3	US-09-041-889-27	Sequence 27, Appl
18	65.5	39.5	214	4	US-09-147-264-27	Sequence 27, Appl
19	65.5	39.5	489	4	US-09-489-039A-13565	Sequence 13565, A
20	65	39.2	223	3	US-09-095-855-201	Sequence 201, App
21	65	39.2	223	4	US-09-205-426-201	Sequence 201, App
22	65	39.2	364	4	US-09-107-532A-5044	Sequence 5044, Ap
23	64.5	38.9	204	4	US-08-529-055-21	Sequence 21, Appl
24	64.5	38.9	585	4	US-08-134-000C-3802	Sequence 3802, Ap
25	64.5	38.9	8991	4	US-08-714-741-32	Sequence 32, Appl
26	64	38.6	700	4	US-09-107-532A-5094	Sequence 5094, Ap
27	63	38.0	148	4	US-09-248-796A-26989	Sequence 26989, A

Query Match 65.1%; Score 108; DB 4; Length 45;
Best Local Similarity 64.4%; Pred. No. 2.4e-06;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY 1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAAAKEAAAYEA 35
||||| ||| ||||| | | : ||||| |||
DB 1 AKKYAKKAAKAAKAYKAAEAKKAAKYKAAEAKAAEAAKEAAAYEA 45
||||| ||| ||||| | | : ||||| |||

RESULT 3
US-09-405-743A-6
; Sequence 6, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-6

Query Match 65.1%; Score 108; DB 4; Length 86;
Best Local Similarity 57.4%; Pred. No. 4.7e-06;
Matches 27; Conservative 2; Mismatches 6; Indels 12; Gaps 1;

QY 1 AKKYAKKEKAACKAY-----KKEAKAKAAEAAAKEAAAYEA 35
||||| ||| ||| | | : ||||| |||
DB 40 AKKYAKAAKAEKKEYAAAEAKYKAAEAKKAYKAAEAKAAEAAAYEA 86
||||| ||| ||| | | : ||||| |||

RESULT 4
US-09-405-743A-4
; Sequence 4, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-4

Query Match 63.6%; Score 105.5; DB 4; Length 66;
Best Local Similarity 43.9%; Pred. No. 7e-06;
Matches 29; Conservative 2; Mismatches 4; Indels 31; Gaps 1;

QY 1 AKKYAKKEKAACKAYKKEA-----KAKAAEAAAK 29
||||| ||| ||| | | : ||||| |||
DB 1 AKKYAKKAYAKAKKAAEAKKAAKAAEAKKAAEAKYKAAEAAKAAAK 60
||||| ||| ||| | | : ||||| |||

QY 30 EAAVEA 35
|||||

DB 61 EAAVEA 66
|||||

RESULT 5
US-09-405-743A-7
; Sequence 7, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-7

Query Match 63.3%; Score 105; DB 4; Length 109;
Best Local Similarity 62.2%; Pred. No. 1.4e-05;
Matches 28; Conservative 2; Mismatches 5; Indels 10; Gaps 2;

QY 1 AKKYAKKEKAACKAY-----KKEA---KAKAAEAAAKEAAAYEA 35
||||| ||| ||| | | : ||||| |||
DB 65 AKKYAKAAKAEKKEYAAAEAKKAAEAKKAYKAAEAKAAEAAAYEA 109
||||| ||| ||| | | : ||||| |||

RESULT 6
US-09-405-743A-3
; Sequence 3, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-3

Query Match 63.0%; Score 104.5; DB 4; Length 56;
Best Local Similarity 51.8%; Pred. No. 7.8e-06;
Matches 29; Conservative 4; Mismatches 2; Indels 21; Gaps 2;

QY 1 AKKYAKKEKA---AKKAYK-----EAKAKAAEAAAKEAAAYEA 35
||||| ||| ||| | | : ||||| |||
DB 1 AKKYAKKAYAKKAAEAKKAAEAKKAAEAKKAYKAAEAKKAAEAAAYEA 56
||||| ||| ||| | | : ||||| |||

RESULT 7
US-09-405-743A-5
; Sequence 5, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1


```

; SEQ ID NO 5
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-5

Query Match          63.0%; Score 104.5; DB 4; Length 77;
Best Local Similarity 65.8%; Pred. No. 1.1e-05;
Matches 25; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 1 AKKYAKKEKAAKKAY---KKEAKAKAAEAAKAAEAYEA 35
Db 40 AKKYAKAAKAEKKEAYAAAEAKYKAEAAKAAKAAEAYEA 77

RESULT 8
US-09-248-796A-18922
; Sequence 18922, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18922
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18922

Query Match          52.1%; Score 86.5; DB 4; Length 176;
Best Local Similarity 65.7%; Pred. No. 0.0034;
Matches 23; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

QY 2 KKYAKKEKAAKKAYKKEAKAKAA-EAAKAAEAYEA 35
Db 84 KEKAKKEKAAKKEKESGSRKAAEAAKAAEAA 118

RESULT 9
US-09-252-991A-29581
; Sequence 29581, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match          45.8%; Score 76; DB 4; Length 407;
Best Local Similarity 56.1%; Pred. No. 0.14;
```

```

Matches 23; Conservative 5; Mismatches 7; Indels 6; Gaps 2;

QY 1 AKKYAK---KEKAAKKAYKK---EAKAKAAEAAKAAEAYEA 35
Db 215 AKKRAEDEAKKAAEDAKKAAEDAKKAAEAAKAAEAAEAA 255

RESULT 10
US-08-152-488-13
; Sequence 13, Application US/08152488
; Patent No. 5534619
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J, Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152.488
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
US-08-152-488-13

Query Match          40.7%; Score 67.5; DB 1; Length 32;
Best Local Similarity 61.3%; Pred. No. 0.097;
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKKYAKK-EKAAKKAYKKEAKAAEAAAKE 30
Db 2 AKKAAKAAKAAKAAKAAKAAKAAKAAKAAK 32

RESULT 11
US-08-303-025-15
```

```

1  NUMBER OF SEQUENCES: 13
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: Benita J, Rohm, Esq.
4  STREET: 512 Springfield Avenue
5  CITY: Cranford
6  STATE: New Jersey
7  COUNTRY: United States of America
8  ZIP: 07016-1811
9
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Floppy disk
12 COMPUTER: IBM PC compatible
13 OPERATING SYSTEM: MS-DOS
14 SOFTWARE: WordPerfect 6; ASCII (DOS)Text
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/677,304
17 FILING DATE:
18 CLASSIFICATION: 530
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER: US/08/152,488
21 FILING DATE: 12-NOV-1993
22 APPLICATION NUMBER: PCT/US92/08069
23 FILING DATE: 14-AUG-1993
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Rohm, Benita J.
26 REGISTRATION NUMBER: 28,664
27 REFERENCE/DOCKET NUMBER: RM-7WG
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: 908-276-3344
30 TELEFAX: 908-276-5543
31 INFORMATION FOR SEQ ID NO: 13:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 32 amino acids
34 TYPE: amino acid
35 STRANDEDNESS: No. 5721212 Relevant
36 TOPOLOGY: No. 5721212 Relevant
37 MOLECULE TYPE: peptide
38 ORIGINAL SOURCE:
39 ORGANISM: N/A
40 PUBLICATION INFORMATION:
41 AUTHORS: N/A
42 TITLE: N/A
43 PUBLICATION INFORMATION:
44 DOCUMENT NUMBER: PCT/US92/08069
45 FILING DATE: 14-AUG-1993
46 US-08-677-304-13
47
48 Query Match 40.7%; Score 67.5; DB
49 Best Local Similarity 61.3%; Pred. No. 0.097
50 Matches 19; Conservative 2; Mismatches
51
52 Qy 1 AKKYAKK-EKAAKAYKKEAKAKAAEAAAKE 30
53 ||||| :||||| :||||| :|||
54 Db 2 AKYAAKKAAYKAKKAAYKAAKKAAYKAKK 32
55 ||||| :||||| :||||| :|||
56
57 RESULT 13
58 US-08-436-703B-2
59 Sequence 2, Application US/08436703B
60 Patent No. 5919761
61 GENERAL INFORMATION:
62 APPLICANT: Wakefield, Thomas W.
63 APPLICANT: Andrews, Philip C.
64 APPLICANT: Stanley, James C.
65 TITLE OF INVENTION: NOVEL PEPTIDES FOR
66 TITLE OF INVENTION: HEPARIN AND LOW MOLECUL
67 TITLE OF INVENTION: WEIGHT HEPARIN
68 TITLE OF INVENTION: ANTICOAGULATION REVERSA
69 NUMBER OF SEQUENCES: 18
70 CORRESPONDENCE ADDRESS:
71 ADDRESSEE: Benita J, Rohm, Esq.
72 STREET: 6601 Woodward Avenue
73 STREET: Suite 1525
74 CITY: Detroit

```

```
/ STATE: Michigan
/ COUNTRY: United States of America
/ ZIP: 48226
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 6;
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/436,703B
/ FILING DATE: 08-MAY-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: N/A
/ FILING DATE: N/A
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rohm, Benita J.
/ REGISTRATION NUMBER: 28,664
/ REFERENCE/DOCKET NUMBER: 7WK-060548-00233
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 313-965-1976
/ TELEFAX: 313-965-1951
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: N/A
/ TOPOLOGY: N/A
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: N/A
/ PUBLICATION INFORMATION:
/ AUTHORS: N/A
/ TITLE: N/A
/ US-08-436-703B-2

Query Match 40.7%; Score 67.5; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 0.097;
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKYAKK-EKAACKAYKKEAKAKAEAAAKE 30
Db 2 AKKAACKAKKAACKAKKAACKAKKAACK 32

RESULT 14
US-08-303-025-16
/ Sequence 16, Application US/08303025
/ Patent No. 561494
/ GENERAL INFORMATION:
/ APPLICANT: Wakefield, Thomas W.
/ APPLICANT: Andrews, Philip C.
/ TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
/ TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Benita J. Rohm, Esq.
/ STREET: 150 West Jefferson, Suite 2500
/ CITY: Detroit
/ STATE: Michigan
/ COUNTRY: United States of America
/ ZIP: 48226-4415
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: MS-DOS v.6.22
/ SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/303,025
/ FILING DATE: 08-SEPT-1994

/ STATE: Michigan
/ COUNTRY: United States of America
/ ZIP: 48226
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 6;
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/436,703B
/ FILING DATE: 08-MAY-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: N/A
/ FILING DATE: N/A
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rohm, Benita J.
/ REGISTRATION NUMBER: 28,664
/ REFERENCE/DOCKET NUMBER: 7WK-060548-00233
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 313-965-1976
/ TELEFAX: 313-965-1951
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 32 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: N/A
/ TOPOLOGY: N/A
/ MOLECULE TYPE: peptide
/ ORIGINAL SOURCE:
/ ORGANISM: N/A
/ PUBLICATION INFORMATION:
/ AUTHORS: N/A
/ TITLE: N/A
/ US-08-436-703B-2

Query Match 40.7%; Score 67.5; DB 1; Length 33;
Best Local Similarity 61.3%; Pred. No. 0.1;
Matches 19; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 AKYAKK-EKAACKAYKKEAKAKAEAAAKE 30
Db 3 AKKAACKAKKAACKAKKAACKAKKAACK 33

RESULT 15
US-08-436-703B-4
/ Sequence 4, Application US/08436703B
/ Patent No. 5919761
/ GENERAL INFORMATION:
/ APPLICANT: Wakefield, Thomas W.
/ APPLICANT: Andrews, Philip C.
/ APPLICANT: Stanley, James C.
/ TITLE OF INVENTION: NOVEL PEPTIDES FOR
/ TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
/ TITLE OF INVENTION: WEIGHT HEPARIN
/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
/ NUMBER OF SEQUENCES: 18
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Benita J. Rohm, Esq.
/ STREET: 6601 Woodward Avenue
/ STREET: Suite 1525
/ CITY: Detroit
/ STATE: Michigan
/ COUNTRY: United States of America
/ ZIP: 48226
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 6;
/ SOFTWARE: ASCII (DOS)Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/436,703B
/ FILING DATE: 08-MAY-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: N/A
/ FILING DATE: N/A
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rohm, Benita J.
```

```

;
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
;
US-08-436-703B-4

Query Match      40.7%; Score 67.5; DB 2; Length 33;
Best Local Similarity 61.3%; Pred. No. 0.1;
Matches 19; Conservative 9; Mismatches 9; Indels 1; Gaps 1;

QY      1 AKYAKK-EKAAKAYKKEAKAAEAAAKE 30
      ||| ||| : ||||| | ||| | |||
Db      3 AKKAAKAKKAAKAAKAAKAAKAAKAAK 33

Search completed: December 14, 2004, 05:50:15
Job time : 10.9635 secs

```


[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 50.5063 Seconds
(without alignments)
512.646 Million cell updates/sec

Title: US-10-792-311-2

Perfect score: 213

Sequence: 1 AKYAKAKAEKAKYKAA.....AKYKAAAEKAAKEAYE 45

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	108.5	50.9	379	2	Q7WFn5	Q7WFn5 bordetella
2	103.5	48.6	210	1	HI_LYTPI	P06144 lytechinus
3	103.5	48.6	373	2	Q7W477	Q7W477 bordetella
4	98.5	46.2	461	2	Q7UR44	Q7UR44 rhodospirillum rubrum
5	98.5	46.0	394	2	Q7AG18	Q7AG18 escherichia coli
6	98.5	46.0	394	2	Q8X965	Q8X965 escherichia coli
7	98.5	46.0	421	1	T0UA_ECOLI	P19934 escherichia coli
8	98.5	46.0	421	2	Q8FJTI	Q8FJTI escherichia coli
9	95.5	44.8	198	2	Q6MIU4	Q6MIU4 bdellovibrio bacteriovorus
10	95.5	44.8	198	2	CAE80819	CAE80819 bdellovibrio bacteriovorus
11	95.5	44.6	413	2	Q7C2Q4	Q7C2Q4 shigella flexneri
12	95.5	44.6	413	2	Q83SA1	Q83SA1 shigella flexneri
13	94.5	44.4	239	2	Q8Y5W4	Q8Y5W4 listeria monocytogenes
14	93.5	43.9	248	2	Q81H69	Q81H69 bacillus subtilis
15	92.5	43.2	117	2	Q89DA0	Q89DA0 bradyrhizobium
16	91.5	42.7	244	2	Q9AJX2	Q9AJX2 streptomyces
17	90.5	42.5	197	2	Q7W3X2	Q7W3X2 bordetella pertussis
18	90.5	42.3	243	2	Q92A67	Q92A67 listeria monocytogenes
19	90.5	42.3	280	2	Q6CEE5	Q6CEE5 yarrowia lipolytica
20	90.5	42.3	1391	1	MS2_DROHY	Q08696 drosophila
21	89.5	42.0	208	1	DBH_MYCSM	Q92HC5 mycobacterium
22	89.5	42.0	211	1	HI8_STRPU	P15869 strongyloides
23	89.5	42.0	214	1	DBH_MYCBO	Q9XB18 mycobacterium
24	89.5	42.0	214	1	DBH_MYCTU	P95109 mycobacterium
25	89.5	42.0	244	2	Q71Y74	Q71Y74 listeria monocytogenes
26	89.5	42.0	244	2	AAT04740	AAT04740 listeria monocytogenes
27	89.5	42.0	549	2	Q75CR3	Q75CR3 ashbya gossypii
28	89.5	42.0	549	2	AAS51084	AAS51084 ashbya gossypii
29	89.5	41.8	182	2	Q8RL36	Q8RL36 burkholderia
30	88.5	41.5	275	1	MS23_DROHY	Q01395 drosophila
31	88.5	41.5	372	2	Q9WXX1	Q9WXX1 pseudomonas

32	88.5	41.5	372	2	Q88NI6	Q88NI6 pseudomonas
33	88	41.3	225	2	Q7WFA2	Q7WFA2 bordetella
34	88	41.3	420	2	Q73D27	Q73D27 bacillus cereus
35	88	41.3	420	2	AAS39818	AAS39818 bacillus cereus
36	88	41.3	899	2	Q9A5J6	Q9A5J6 caulobacter
37	87.5	41.1	1551	2	Q7SBU1	Q7SBU1 neurospora
38	86.5	40.6	200	2	Q8XVW7	Q8XVW7 ralstonia
39	86.5	40.6	384	1	TMPB_TREPH	P29720 treponema pallidum
40	86	40.4	105	2	QGN503	Q6N503 rhodospirillum rubrum
41	86	40.4	105	2	CAE28621	CAE28621 rhodospirillum rubrum
42	86	40.4	155	2	Q8PI40	Q8PI40 xanthomonas
43	86	40.4	168	1	RS16_COREF	Q8PI30 corynebacterium
44	86	40.4	198	1	HBHA_MYCTU	Q11142 mycobacterium
45	86	40.4	212	2	O93946	O93946 candida albicans

ALIGNMENTS

RESULT 1

ID	Q7WFn5	PRELIMINARY;	PRT;	379 AA.
AC	Q7WFn5;			
DT	01-OCT-2003 (TRENBLrel. 25, Created)			
DT	01-OCT-2003 (TRENBLrel. 25, Last sequence update)			
DE	01-MAR-2004 (TRENBLrel. 26, Last annotation update)			
GN	OrderedLocusNames=BB4236;			
OS	Bordetella bronchiseptica (Alcaligenes bronchisepticus).			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC	Alcaligenaceae; Bordetella.			
OX	NCBI_TaxID=518;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=RB50 / ATCC BAA-588;			
RX	MEDLINE=2921954; PubMed=12910271; DOI=10.1038/ngi227;			
RA	Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,			
RA	Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,			
RA	Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,			
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,			
RA	Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,			
RA	Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,			
RA	Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,			
RA	Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,			
RA	Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,			
RA	Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;			
RT	"Comparative analysis of the genome sequences of Bordetella pertussis,			
RT	Bordetella parapertussis and Bordetella bronchiseptica.";			
RL	Nat. Genet. 35:32-40(2003).			
DR	EMBL; BX640449; CAE34600.1; -			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.			
DR	GO; GO:0008565; F:protein transporter activity; IEA.			
DR	GO; GO:0015031; P:protein transport; IEA.			
DR	InterPro; IPR010528; T01A.			
DR	InterPro; IPR006260; TonB_C.			
DR	Pfam; PF06519; T01A; 1.			
DR	TIGRFAMs; TIGR01352; tonB_C-term; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 379 AA; 40776 MW; C657B5AAE97EBDD3 CRC64;			

Query Match 50.9%; Score 108.5; DB 2; Length 379;
Best Local Similarity 65.1%; Pred. No. 0.0073;
Matches 28; Conservative 5; Mismatches 7; Indels 3; Gaps 1;

QY 5 AKKAAEKAKYKAAEKAAKYE---KAAEKAAEKAAEAYE 44

DB 201 AKKAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAE 243

RESULT 2

HI_LYTPI STANDARD; PRT; 210 AA.

AC P06144;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Late histone H1.
OS Lytechinus pictus (Painted sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7653;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX Knowles J.A., Childs G.J.;
RA "Comparison of the late H1 histone genes of the sea urchins Lytechinus
RT pictus and Strongylocentrotus purpuratus";
RL Nucleic Acids Res. 14:18121-18133(1986).
CC 1- FUNCTION: Histones H1 are necessary for the condensation of
CC nucleosome chains into higher order structures.
CC 1- SUBCELLULAR LOCATION: Nuclear.
CC 1- SIMILARITY: Belongs to the histone H1/H5 family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04488; CAA28177.1; -.
DR PIR; A25550; A25550.
DR HSSP; P02259; IHST.
DR InterPro; IPR005818; Histone_H1/H5.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR003216; Linkerhist_N.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00538; Linker_histone; 1.
DR PRINTS; PR00624; HISTONER5.
DR ProDom; PD000373; Linkerhist_N; 1.
DR SMART; SM00526; H15; 1.
DR Chromosomal protein; DNA-binding; Multigene family; Nuclear protein.
KW SEQUENCE 210 AA; 21746 MW; 08C38F64894007E2 CRC64;
SQ
Query Match 48.6%; Score 103.5; DB 1; Length 210;
Best Local Similarity 65.1%; Pred. No. 0.013;
Matches 28; Conservative 3; Mismatches 11; Indels 1; Gaps 1;
QY 1 AKKYAKKAAEK-AKKAYKAAEAKKAAKYEKAAAEKAAAEKAA 42
DB 166 AKKAAKPAKAKPAKAAKPAKAAKPAKAAKPAKAAKPAKAA 208
RESULT 3
QY477 ID QY477 PRELIMINARY; PRT; 373 AA.
AC QY477;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proline-rich inner membrane protein.
GN OrderedLocustNames=BPP3791;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,

RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640434; CAE39074.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0015033; P:protein transport; IEA.
DR InterPro; IPR010528; ToIA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; ToIA; 1.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 373 AA; 40092 MW; A364894DAE46E328 CRC64;
Query Match 48.6%; Score 103.5; DB 2; Length 373;
Best Local Similarity 62.7%; Pred. No. 0.022;
Matches 32; Conservative 4; Mismatches 8; Indels 7; Gaps 3;
QY 1 AKKYA--KKAAEKAKK--AYKAAEAKKAAKYE---KAAAEKAAAEKAAE 44
DB 187 AEKQAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAEKAAE 237
RESULT 4
QY477 ID QY477 PRELIMINARY; PRT; 461 AA.
AC QY477;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=RH5786;
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]_TaxID=117;
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294143; CAD74436.1; -.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 7.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PS50005; TPR; 4.
DR PROSITE; PS50293; TPR REGION; 1.
KW Complete proteome; Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 461 AA; 50441 MW; A7C5752864CBE14C CRC64;
Query Match 46.2%; Score 98.5; DB 2; Length 461;
Best Local Similarity 57.8%; Pred. No. 0.084;
Matches 26; Conservative 5; Mismatches 13; Indels 1; Gaps 1;
QY 2 KKY-AKKAAEKAKKAAEKAAEKAAEKAAEKAAEKAAEKAAE 45
DB 404 KPYIADPAAAEKAAEAAABIEAKKAAEKAAEKAAEKAAEKAAE 448

```

RESULT 5
ID Q7AGI8 PRELIMINARY; PRT; 394 AA.
AC Q7AGI8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Membrane spanning protein TolA.
OS OrderedLocusNames=ECs0774;
GN Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
DR EMBL; AP002553; BAB34197.1; -.
DR InterPro; IPR010528; TolA.
DR Pfam; PF06519; TolA; 1.
SQ SEQUENCE 394 AA; 40517 MW; 5B58DB8B230BDE28 CRC64;

Query Match 46.0%; Score 98; DB 2; Length 394;
Best Local Similarity 65.3%; Pred. NO. 0.081;
Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

OY 2 KKYAKKAKEA---KK--AYKAAEAKKAAYEKAAAEKAAAEKAAAEKAAAE 45
DB 220 KKAEEKAAAEKAAADKKAEEKAAADKKAEEKAAAEKAAAEKAAAEKAAAE 267

RESULT 6
ID Q8X965 PRELIMINARY; PRT; 394 AA.
AC Q8X965;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 26, Last annotation update)
DE Membrane spanning protein, required for outer membrane integrity.
GN Name=tolA; OrderedLocusNames=z0907;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouisis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR EMBL; AE005252; AAG55075.1; -.
DR PIR; F90725; F90725.
DR PIR; G85576; G85576.
DR HSSP; PI9934; ITOL.
DR InterPro; IPR010528; TolA.
DR Pfam; PF06519; TolA; 1.
KW Complete proteome.
SQ SEQUENCE 394 AA; 40517 MW; 5B58DB8B230BDE28 CRC64;

Query Match 46.0%; Score 98; DB 2; Length 394;
Best Local Similarity 65.3%; Pred. NO. 0.081;
Matches 32; Conservative 2; Mismatches 9; Indels 6; Gaps 3;

OY 2 KKYAKKAKEA---KK--AYKAAEAKKAAYEKAAAEKAAAEKAAAEKAAAE 45
DB 220 KKAEEKAAAEKAAADKKAEEKAAADKKAEEKAAAEKAAAEKAAAEKAAAE 267

RESULT 7
TOL A_ECOLI STANDARD; PRT; 421 AA.
AC P19334;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE TolA protein.
GN Name=tolA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levgood S.K., Webster R.E.;
RT "Nucleotide sequences of the tolA and tolB genes and localization of
RT their products, components of a multistep translocation system in
RT Escherichia coli.";
RL J. Bacteriol. 171:6600-6609(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levgood S.K., Beyer W.F. Jr., Webster R.E.;
RT "TolA: a membrane protein involved in colicin uptake contains an
RT extended helical region.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RN [5]
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "TolA central domain interacts with Escherichia coli porins.";
RL EMBO J. 15:6408-6415(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX MEDLINE=99332679; PubMed=10404600;
RA Lubkowski J., Hennecke F., Plueckthun A., Wlodawer A.;
RT "Filamentous phage infection: crystal structure of g3p in complex with
RT its coreceptor, the C-terminal domain of TolA.";

```


Query Match 44.6%; Score 95; DB 2; Length 413;
Best Local Similarity 63.3%;
Pred. No. 0.17;
Matches 31; Conservative 3; Mismatches 9; Indels 6; Gaps 3;

SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P, Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Debou P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etienne K.-D., Feihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kraft J., Kuhn M., Kunst F., Kurapkat G.,
RA Maqueno E., Maicournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 49.7468 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-2
Perfect score: 213
Sequence: 1 AKYAKAKAEKAKYKAA.....AKYKAAAEKAAKEAYEA 45

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	213	100.0	45	3	AAy82572 Copolymer
2	138	64.8	109	3	AAy82577 Copolymer
3	134.5	63.1	56	3	AAy82573 Copolymer
4	131	61.5	77	3	AAy82575 Copolymer
5	126.5	59.4	86	3	AAy82576 Copolymer
6	120.5	56.6	66	3	AAy82574 Copolymer
7	108.5	50.9	428	6	ABU27824 Protein e
8	108	50.7	35	3	AAy82571 Copolymer
9	99.5	46.7	106	2	Aar06446 Recombina
10	98	46.0	46	2	Aar28871 High affi
11	98	46.0	421	6	ABU28559 Protein e
12	95	44.6	323	6	ABU31397 Protein e
13	95	44.6	469	7	ABO67048 Klebsiell
14	94.5	44.4	239	5	ABb49123 Listeria
15	94.5	44.4	239	6	ABU32619 Protein e
16	94	44.1	154	2	Aar06445 Recombina
17	91	42.7	223	2	AAy14928 Amino aci
18	91	42.7	223	6	ABP70903 Mycobacte
19	89.5	42.0	214	2	AAy34055 M. tuberc
20	89.5	42.0	214	2	AAy57353 M. tuberc
21	89.5	42.0	214	6	ABU34623 Protein e
22	89.5	42.0	214	6	ABU36893 Protein e
23	88.5	41.5	372	6	ABU40185 Protein e
24	87	40.8	334	4	ABG28693 Novel hum
25	86	40.4	39	2	AAW44934 Mycobacte

26	86	40.4	39	7	ADP45561	Adf45561 Methylate
27	86	40.4	198	2	AAW43082	Aaw43082 Mycobacte
28	86	40.4	198	2	AAW44936	Aaw44936 Mycobacte
29	86	40.4	347	6	ABU38313	Abu38313 Protein e
30	86	40.4	347	6	ABJ18771	Abj18771 Pseudomon
31	86	40.4	407	7	ABO80835	AbO80835 Pseudomon
32	85.5	40.1	205	3	AAE20575	AAE20575 Mycobacte
33	85	39.9	427	4	AAg70868	AAg70868 C albican
34	85	39.9	427	8	ADP99051	ADp99051 C. albica
35	83	39.0	389	6	ABU39221	ABu39221 Protein e
36	82	38.5	165	4	AAg91997	AAg91997 C glutami
37	81.5	38.3	357	6	ABM67869	ABm67869 Photorhab
38	81.5	38.3	497	7	ADH88105	ADh88105 Enterococ
39	81	38.0	80	5	ABG71044	ABg71044 Tumour ne
40	79.5	37.3	407	6	ABU47123	ABu47123 Protein e
41	79	37.1	212	6	ABU33811	ABu33811 Protein e
42	78.5	36.9	875	6	ABU22879	ABu22879 Protein e
43	78	36.6	157	6	ABU26064	ABu26064 Protein e
44	78	36.6	309	7	ABO23523	ABo23523 Pseudomon
45	78	36.6	316	7	ABO84211	ABo84211 Pseudomon

ALIGNMENTS

RESULT 1
AAy82572
ID AAy82572 standard; peptide; 45 AA.

XX AAy82572;

XX 28-JUL-2000 (first entry)

XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:2.

XX Copolymer; molecular weight marker; TV-marker; immune disease;
XX glairamer acetate; autoimmune disease; antithyroid; neuroprotective;
XX osteopathic; immunosuppressive; antithyroid; antiinflammatory;
XX antididiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
XX antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
XX inflammatory condition; multiple sclerosis; rheumatoid arthritis;
XX Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
XX diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
XX Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
XX pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA) YEDA RES & DEV CO LTD.
(TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for
glairamer acetate and for treatment and prevention of immune diseases.

XX Claim 10; Page 14; 72pp; English.

XX AAy82571 to AAy82577 represent specifically claimed copolymer molecular
weight TV-marker polypeptides from the present invention. The present
invention describes polypeptides (I) for determining the molecular weight
of a copolymer (CP), which has an identified molecular weight and an
amino acid composition corresponding to the copolymer. The polypeptides

CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 45 AA;

Query Match 100.0%; Score 213; DB 3; Length 45;
 Best Local Similarity 100.0%; Pred. No. 3.2e-16;
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKKYAKKAKAEKAKYKAAEAKKAAKYEKAAAEKAAAEKAAEAYEA 45
 |||||
 DB 1 AKKYAKKAKAEKAKYKAAEAKKAAKYEKAAAEKAAAEKAAEAYEA 45

RESULT 2
 AAY82577
 ID AAY82577 standard; peptide; 109 AA.

AC AAY82577;
 XX
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:7.
 XX
 KW Copolymer; molecular weight marker; TV-marker; immune disease;
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.
 XX
 XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA) YEDA RES & DEV CO LTD.
 XX (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for
 XX glatiramer acetate and for treatment and prevention of immune diseases.

PS Claim 10; Page 14; 72pp; English.

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular

CC weight TV-marker polypeptides from the present invention. The present
 CC invention describes polypeptides (I) for determining the molecular weight
 CC of a copolymer (CP) which has an identified molecular weight and an
 CC amino acid composition corresponding to the copolymer. The polypeptides
 CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 109 AA;

Query Match 64.8%; Score 138; DB 3; Length 109;
 Best Local Similarity 76.6%; Pred. No. 1.1e-07;
 Matches 36; Conservative 0; Mismatches 7; Indels 4; Gaps 2;

QY 1 AKKYAKKAKAEKAKYKAAEAKK--AAKYEKAAAEKAAAEKAAEAYEA 45
 |||||
 DB 65 AKKYAKKAAE--KKEYAAAEAKKAAEAKYKAAAEKAAAEAYEA 109

RESULT 3
 AAY82573
 ID AAY82573 standard; peptide; 56 AA.

AC AAY82573;

XX 28-JUL-2000 (first entry)

XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:3.

XX Copolymer; molecular weight marker; TV-marker; immune disease;
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA) YEDA RES & DEV CO LTD.
 XX (TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for
 XX glatiramer acetate and for treatment and prevention of immune diseases.

XX PS Claim 10; Page 14; 72pp; English.

XX CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular

XX CC weight TV-marker polypeptides from the present invention. The present

XX CC invention describes polypeptides (I) for determining the molecular weight

XX CC of a copolymer (CP), which has an identified molecular weight and an

XX CC amino acid composition corresponding to the copolymer. The polypeptides

XX CC of the invention are used as molecular weight markers for glatiramer

XX CC acetate related tetrapolymers. The polypeptides may also be used for

XX CC treating and preventing immune diseases in a mammal. Autoimmune diseases

XX CC which may be treated include either cell-mediated or antibody-mediated

XX CC diseases. Such diseases include arthritic conditions, demyelinating

XX CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune

XX CC disease, chronic immune thrombocytopaenia purpura, colitis, contact

XX CC disease, chronic immune thyroiditis, autoimmune uveoretinitis, Crohn's

XX CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's

XX CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,

XX CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-

XX CC mediated diseases which can be treated include host-versus-graft disease,

XX CC graft-versus-host disease, and delayed-type hypersensitivity. The

XX CC polypeptides of the invention have defined molecular weights and physical

XX CC properties which are analogous to glatiramer acetate molecules, which

XX CC makes them ideal for use as molecular weight markers

XX CC Sequence 56 AA;

XX CC

XX CC Query Match 63.1%; Score 134.5; DB 3; Length 56;

XX CC Best Local Similarity 68.4%; Pred. No. 1.4e-07;

XX CC Matches 39; Conservative 0; Mismatches 5; Indels 13; Gaps 4;

XX CC

QY 1 AKKYAKK-----AKAEKA-----KKAYKAAEAKK--AAKYKAAAEKAAKEAAVEA 45

DB 1 AKKYAKKEKAYAKKAEKAAKAEKAAKAYKAAEAKKAEAKY-KAEAKAAKAAKEAAVEA 56

RESULT 4

AAV82575

ID AAY82575 standard; peptide; 77 AA.

AC AAY82575;

XX 28-JUL-2000 (first entry)

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:5.

XX Copolymer; molecular weight marker; TV-marker; immune disease;

KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;

KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;

KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;

KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;

KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;

KW Hashimoto's disease; Graves disease; Guillain-Barre's syndrome; psoriasis;

KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

OS

XX WO200018794-A1.

PN

XX 06-APR-2000.

PD

XX 24-SEP-1999; 99WO-US022402.

PF

XX 25-SEP-1998; 98US-0101693P.

PR

XX (YEDA) YEDA RES & DEV CO LTD.

PA (TEVA-) TEVA PHARM USA INC.

PA

XX Gad A, Lis D;

PI

DR WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for

PT glatiramer acetate and for treatment and prevention of immune diseases.

PT

XX Claim 10; Page 14; 72pp; English.

PS

XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular

CC weight TV-marker polypeptides from the present invention. The present

CC invention describes polypeptides (I) for determining the molecular weight

CC of a copolymer (CP), which has an identified molecular weight and an

CC amino acid composition corresponding to the copolymer. The polypeptides

CC of the invention are used as molecular weight markers for glatiramer

CC acetate related tetrapolymers. The polypeptides may also be used for

CC treating and preventing immune diseases in a mammal. Autoimmune diseases

CC which may be treated include either cell-mediated or antibody-mediated

CC diseases. Such diseases include arthritic conditions, demyelinating

CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid

CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune

CC disease, chronic immune thyroiditis, autoimmune uveoretinitis, Crohn's

CC disease, chronic immune thrombocytopaenia purpura, colitis, contact

CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's

CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,

CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-

CC mediated diseases which can be treated include host-versus-graft disease,

CC graft-versus-host disease, and delayed-type hypersensitivity. The

CC polypeptides of the invention have defined molecular weights and physical

CC properties which are analogous to glatiramer acetate molecules, which

CC makes them ideal for use as molecular weight markers

XX CC Sequence 77 AA;

XX CC

XX CC Query Match 61.5%; Score 131; DB 3; Length 77;

XX CC Best Local Similarity 50.6%; Pred. No. 4.6e-07;

XX CC Matches 39; Conservative 0; Mismatches 6; Indels 32; Gaps 3;

XX CC

QY 1 AKKYAKK-----AKAEKA-----KKAYKAAEAKKAAKYB----- 29

DB 1 AKKYAKKEKAYAKKAEKAAKAEKAAKAYKAAEAKKAEAKKAYKAAAEKKEKAAEAK 60

QY 30 -KAAAEKAAKEAAVEA 45

DB 61 YKAAEAKAAKEAAVEA 77

RESULT 5

AAV82576

ID AAY82576 standard; peptide; 86 AA.

XX AAY82576;

AC

XX 28-JUL-2000 (first entry)

DT

XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.

DE

XX Copolymer; molecular weight marker; TV-marker; immune disease;

KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;

KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;

KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;

KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;

KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;

KW Hashimoto's disease; Graves disease; Guillain-Barre's syndrome; psoriasis;

KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;

KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

OS

XX WO200018794-A1.

PN

XX 06-APR-2000.

PD

XX 24-SEP-1999; 99WO-US022402.

PF

PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 PA (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA35267.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 59321; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 323 AA;
 SQ
 Query Match 44.6%; Score 95; DB 6; Length 323;
 Best Local Similarity 52.9%; Pred. No. 0.016;
 Matches 27; Conservative 4; Mismatches 14; Indels 6; Gaps 1;
 QY 1 AKKYAKKAAEKAKYKAAEAK-----KAAKYKAAAEKAAKAAEAAVEA 45
 Db 92 ABAEQKAAEAAAKKAAQAEKKAQAEAKQAQAEAKAAEAAEAAEAAQKA 142
 RESULT 13
 ABO67048
 ID ABO67048 standard; protein; 469 AA.
 XX
 AC ABO67048;
 XX
 XX 29-JUL-2004 (first entry)
 DT
 XX Klebsiella pneumoniae polypeptide seqid 13565.
 DE
 XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.
 OS
 XX US6610836-B1.
 PN
 XX 26-AUG-2003.
 PD
 XX 27-JAN-2000; 2000US-00489039.
 PF
 XX 29-JAN-1999; 99US-0117747P.
 PR
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Breton GL, Osborne M;
 PI
 XX WPI; 2003-895346/82.
 DR N-PSDB; ABD00619.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 PT
 XX Disclosure; SEQ ID NO 13565; 932pp; English.
 PS
 XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 XX Sequence 469 AA;
 SQ
 Query Match 44.6%; Score 95; DB 7; Length 469;
 Best Local Similarity 52.9%; Pred. No. 0.024;
 Matches 27; Conservative 4; Mismatches 14; Indels 6; Gaps 1;
 QY 1 AKKYAKKAAEKAKYKAAEAK-----KAAKYKAAAEKAAKAAEAAVEA 45
 Db 242 ABAEQKAAEAAAKKAAQAEKKAQAEAKQAQAEAKAAEAAEAAEAAQKA 292
 RESULT 14
 ABB49123
 ID ABB49123 standard; protein; 239 AA.
 XX
 AC ABB49123;
 XX
 DT 05-FEB-2002 (first entry)
 DT
 XX Listeria monocytogenes protein #1827.
 DE
 XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 KW
 XX Listeria monocytogenes.
 OS
 XX WO200177335-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX 11-APR-2001; 2001WO-FR001118.
 PF
 XX 11-APR-2000; 2000FR-00004629.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihi H, Dehoux P;
 PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Pierrez-Martinez A, Amend A;
 PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

This Page Blank (uspto)

RESULT 2


```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match          50.9%; Score 108.5; DB 15; Length 428;
Best Local Similarity 60.4%; Pred. No. 0.00079;
Matches 32; Conservative 2; Mismatches 8; Indels 11; Gaps 2;

QY      1 AKKYA---KKAAEKAKKAYKAAE-----AKKAAKYEKAAAEKAAAEAA 42
Db      199 AKKAAADAQKAAEAEEKAAQAAEKAAAEAAKAAAEKAAAEKAAAEKAA 251

RESULT 8
US-09-816-989A-1
; Sequence 1, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WE
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-1

Query Match          50.7%; Score 108; DB 9; Length 35;
Best Local Similarity 64.9%; Pred. No. 6.1e-05;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY      1 AKKYAKKAAEKAKKAYKAAEAEEKAAKYEKAAAEKAAAEKAAAEAA 45
Db      1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAEEKAAAEAA 35

RESULT 9
US-10-282-122A-56483
; Sequence 56483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

```



```
; SEQ ID NO 60543
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60543

Query Match          44.4%; Score 94.5; DB 15; Length 239;
Best Local Similarity 56.2%; Pred. No. 0.014;
Matches 27; Conservative 7; Mismatches 9; Indels 5; Gaps 2;

QY 2 KKYAKKAKAEKAKK--AYKAAEAKK--AAKYEKAAAEKAAAEKAAE 44
Db 124 KAAAEKAEADKKQBEDAVKAAANAKKEQEAEEKAAADKAAAEKAAAE 171

RESULT 12
US-10-051-643-201
; Sequence 201, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.11008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 201
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-051-643-201

Query Match          42.7%; Score 91; DB 13; Length 223;
Best Local Similarity 59.6%; Pred. No. 0.03;
Matches 28; Conservative 5; Mismatches 12; Indels 2; Gaps 2;

QY 1 AKKYA-KKAKAEKAKKAYKAAAEKAAKAEKAA-AEKAAAEKAAE 45
Db 137 AKKAATKAPAKKATAAKKAAKAPAKKATAAKKAPAKKAPAKKATA 183

RESULT 13
US-10-205-979-52
; Sequence 52, Application US/10205979
; Publication No. US20030147861A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Abernethy, Nevin
; TITLE OF INVENTION: Compounds and Methods for the Modulation
; TITLE OF INVENTION: of Immune Responses
; FILE REFERENCE: 11000.10630
; CURRENT APPLICATION NUMBER: US/10/205,979
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/308,446
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-205-979-52

Query Match          42.7%; Score 91; DB 14; Length 223;
```

```
Best Local Similarity 59.6%; Pred. No. 0.03;
Matches 28; Conservative 5; Mismatches 12; Indels 2; Gaps 2;

QY 1 AKKYA-KKAKAEKAKKAYKAAAEKAAKAEKAA-AEKAAAEKAAE 45
Db 137 AKKAATKAPAKKATAAKKAAKAPAKKATAAKKAPAKKAPAKKATA 183

RESULT 14
US-10-229-567-27
; Sequence 27, Application US/10229567
; Publication No. US20030092080A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/10/229,567
; APPLICATION NUMBER: US-10/229,567
; FILING DATE: 27-Aug-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/417,264
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 09/041,889
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-10-229-567-27

Query Match          42.0%; Score 89.5; DB 14; Length 214;
Best Local Similarity 55.6%; Pred. No. 0.042;
Matches 25; Conservative 4; Mismatches 13; Indels 3; Gaps 1;

QY 1 AKKYAKKAKAEKAKKAYKAAAEKAAKAEKAAAEKAAAEKAAE 45
Db 111 AKKVAKKAPAKKATKAAKKAATKAPA---RKAATKAPAKKAATKA 152

RESULT 15
US-10-282-122A-62547
; Sequence 62547, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
```

```

/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Onisen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 62547
/ LENGTH: 214
/ TYPE: PR1
/ ORGANISM: Mycobacterium bovis
US-10-282-122A-62547

```

```

Query Match      42.0%; Score 89.5; DB 15; Length 214;
Best Local Similarity 55.6%; Pwd.No. 0.042;
Matches 25; Conservative 4; Mismatches 13; Indels 3; Gaps 1;

QY 1 AKKYAKKAAKAEKAKYAKAAEAKKAAKYEKAAAEKAAAEKAAEAA 45
    |||||
Db 111 AKKVAKKAPAKKATKAAPKAAKAAATKAPA---RKAATKAPAKKAATKA 152
    |||||

```

Search completed: December 14, 2004, 06:59:40
Job time : 111.646 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	213	100.0	45	4	US-09-405-743A-2	Sequence 2, Appli
2	138	64.8	109	4	US-09-405-743A-7	Sequence 7, Appli
3	134.5	63.1	56	4	US-09-405-743A-3	Sequence 3, Appli
4	131	61.5	77	4	US-09-405-743A-5	Sequence 5, Appli
5	126.5	59.4	86	4	US-09-405-743A-6	Sequence 6, Appli
6	120.5	56.6	66	4	US-09-405-743A-4	Sequence 4, Appli
7	108	50.7	35	4	US-09-405-743A-1	Sequence 1, Appli
8	95	44.6	469	4	US-09-489-039A-13565	Sequence 13565, A
9	91	42.7	223	3	US-09-095-855-201	Sequence 201, App
10	91	42.7	223	3	US-09-205-426-201	Sequence 201, App
11	89.5	42.0	214	3	US-09-041-889-27	Sequence 27, Appl
12	89.5	42.0	214	3	US-09-417-264-27	Sequence 27, Appl
13	86	40.4	148	4	US-09-248-796A-26989	Sequence 26989, A
14	86	40.4	407	4	US-09-252-991A-23581	Sequence 23581, A
15	85	39.9	176	4	US-09-248-796A-18922	Sequence 18922, A
16	81.5	38.3	497	4	US-09-134-000C-5390	Sequence 5390, Ap
17	78	36.6	316	4	US-09-252-991A-32957	Sequence 32957, A
18	77.5	36.4	700	4	US-09-107-532A-5094	Sequence 5094, Ap
19	77	36.2	103	3	US-09-041-889-39	Sequence 39, Appl
20	77	36.2	103	4	US-09-417-264-39	Sequence 39, Appl
21	77	36.2	116	3	US-09-041-889-38	Sequence 38, Appl
22	77	36.2	116	4	US-09-417-264-38	Sequence 38, Appl
23	77	36.2	158	3	US-09-041-889-40	Sequence 40, Appl
24	77	36.2	158	4	US-09-417-264-40	Sequence 40, Appl
25	77	36.2	222	3	US-09-041-889-3	Sequence 3, Appli
26	77	36.2	222	3	US-08-837-058-3	Sequence 3, Appli
27	77	36.2	222	4	US-09-417-264-3	Sequence 3, Appli


```

; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-1

Query Match          50.7%; Score 108; DB 4; Length 35;
Best Local Similarity 64.4%; Pred. No. 2.2e-05;
Matches 29; Conservative 1; Mismatches 5; Indels 10; Gaps 2;

QY      1 AKKYAKKAAKKAYKAAAEKKAAYEKAAKAAEKEAAEAEEA 45
||||| ||| ||||| | | | | | | | | | | | | | | | |
Db      1 AKKYAKKEKA--AKKAYK-----KEAKAKAAEAAAEEA 35

RESULT 8
US-09-489-039A-13565
; Sequence 13565, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13565
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13565

Query Match          44.6%; Score 95; DB 4; Length 469;
Best Local Similarity 52.9%; Pred. No. 0.0076;
Matches 27; Conservative 4; Mismatches 14; Indels 6; Gaps 1;

QY      1 AKKYAKKAAKKAKAYKAAEAK-----KAAKYEKAAAEKAAEAAVEA 45
| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      242 AAEAQKAAEAAAKAAQQEAEKKAQQEAAKQAAAEKAAEAAEAAQA 292

RESULT 9
US-09-095-855-201
; Sequence 201, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

Db 137 AKKAATKAAPAKKATAAKKAAPAKKATAAKKAAPAKKATAAKKAATKA 183

RESULT 11

US-09-041-889-27
; Sequence 27, Application US/09041889
; Patent No. 6033864
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Conavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC pANCA antigens
; TITLE OF INVENTION: Microbial UC pANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:

ADDRESS: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA

```

:
:
: ZIP: 92122
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS
:
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:
: APPLICATION NUMBER: US/09/041,889
:
: FILING DATE:
:
:

```

```

1 CLASSIFICATION:
2
3 PRIOR APPLICATION DATA:
4
5 APPLICATION NUMBER: US 08/837,058
6
7 FILING DATE: 11-APR-1997
8
9 ATTORNEY/AGENT INFORMATION:
10
11 NAME: Campbell, Cathryn A.
12
13 REGISTRATION NUMBER: 31,815
14
15 REFERENCE/DOCKET NUMBER: P-PM 3006
16
17 TELECOMMUNICATION INFORMATION:
18
19 TELEPHONE: (619) 535-9001
20
21 TELEFAX: (619) 535-8949
22
23 INFORMATION FOR SEQ ID NO: 27:
24
25 SEQUENCE CHARACTERISTICS:
26
27 LENGTH: 214 amino acids
28
29 TYPE: amino acid
30
31 TOPOLOGY: linear
32
33 MOLECULE TYPE: peptide
34
35 US-09-041-889-27

```

Query Match	42.0%	Score 89.5;	DB 3;	Length 214;
Best Local Similarity	55.6%	Pred. NO. 0.013;		
Matches 25;	Conservative	4;	Mismatches 13;	Indels 3;
Gaps 1;				

Qy 1 AKYAKKAKAEKAKKAYKAAEAKKAKYKAAAEKAAAEKAAAEYA 45
 111 AKTVAKKAPAKKATKAAKKAATKAPAAA:|||||:|
 Db 111 AKTVAKKAPAKKATKAAKKAATKAPAAA--RKAATKAPAKKAATKA 152

RESULT 12

US-09-417-264--27
; Sequence 27, Application US/09417264
; Patent No. 6537768
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC pANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California

```
,  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/417,264  
FILING DATE:
```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/041,889
FILING DATE:

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-417-264-27

```

Query Match	42.0%;	Score 89.5;	DB 4;	Length 214;
Best Local Similarity	55.6%;	Pred. No. 0.013;		
Matches 25;	Conservative	4;	Mismatches 13;	Indels 3;
				Gaps 1;

Qy 1 AKKYAKKAEKAKKAYKAAEAKKAYEKAAAAEKAAYEA 45
 ||| ||| : ||| | : ||| : ||| : ||| :
 Db 111 AKKYAKKAPAKKATKAAKKAATKAPA---RKAATKAPAKKAATKA 152

RESULT 13

```

US-09-248-796A-26989
; Sequence 26989, Application US/09248796A
; Patent No. 674137
;
; GENERAL INFORMATION:
;
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26989
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26989

```

Query Match 40.4%; Score 86; DB 4; Length 148;
Best Local Similarity 48.9%; Pred. No. 0.021;
Matches 23; Conservative 10; Mismatches 12; Indels 2; Gaps 1;

Qy

1 AKKVAKKA--KAEKAKAYKAAEAKKAAYEKAAAEKAAAEEAAYEA 45
||| :| | : ||| : : |:| : ||
D6 6 AKKAAEEAKKKEEEKAKAAEEAKKVEEAAKKAAEEAKKAEAEA 52
||| :| | : ||| : : |:| : ||

RESULT 14

US-09-252-991A-29581
; Sequence 29581, Application US/09252991A
; Patent No. 6551795

```
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 40.4%; Score 86; DB 4; Length 407;
Best Local Similarity 51.1%; Pred. NO. 0.06;
Matches 24; Conservative 6; Mismatches 15; Indels 2; Gaps 1;

QY 1 AKKYAKKAAKAKKA--YKAAEAKKAAKYEKAAAEKAAKAAEAAYA 45
DB 185 AOKAAEAKKAAEAKKAAEAKKAAEQKQADIAKKRAEAEKKAEDA 231

RESULT 15
US-09-248-796A-18922
; Sequence 18922, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18922
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18922

Query Match 39.9%; Score 85; DB 4; Length 176;
Best Local Similarity 50.0%; Pred. NO. 0.033;
Matches 27; Conservative 7; Mismatches 8; Indels 12; Gaps 3;

QY 2 KKYAKKAAKAKKAY-----KAAE---AKKAAKYEKAAAEKAAKAAEAAYA 45
DB 81 KLFKEKAKKEKAKWEKSGSRKAAEEAAKAA--EEAAAKKAAEEAAALKA 132
```

Search completed: December 14, 2004, 05:50:16
Job time : 13.7694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 12.8776 Seconds
(without alignments)
418.411 Million cell updates/sec

```

Title: US-10-792-311-3
Perfect score: 266
Sequence: 1 AKKYANKEAYAKAEKAAK.....EAKYKAERAAKAAKEAAVEA 56

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database :      PIR_79:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	117	44.0	421	2	JV0057	tolA protein - Esc	
2	116.5	43.8	239	2	AE1317	hypothetical prote	
3	113	42.5	334	2	F90725	membrane spanning	
4	113	42.5	394	2	G85576	membrane spanning	
5	110.5	41.5	376	2	AG0592	tolA protein [impo	
6	110	41.4	1701	2	T09127	probable erythrocy	
7	105.5	39.7	372	2	G64064	outer membrane int	
8	102.5	38.5	210	2	A25550	histone H1 - sea u	
9	102.5	38.5	347	2	E83525	TolA protein PA097	
10	102	38.3	211	2	A28100	histone H1-beta, e	
11	100.5	37.8	1390	2	S51364	sperm tail-specifi	
12	99.5	37.4	206	1	HSTR1R	histone H1 - rainb	
13	98.5	37.0	243	2	AE1689	hypothetical prote	
14	98.5	37.0	311	2	T17598	hypothetical prote	
15	97.5	36.7	206	2	S09388	histone H1 - sea u	
16	97	36.5	219	2	E60110	repetitive protein	
17	96	36.1	384	2	B43592	outer membrane pro	
18	95.5	35.9	217	2	A26721	histone H1-gamma,	
19	95.5	35.9	248	1	HSUR1P	histone H1, gonada	
20	95.5	35.9	1403	2	T11583	probable translati	
21	95	35.7	388	2	AC0138	TolA colicin impor	
22	94.5	35.5	214	2	G70673	probable hupB - My	
23	94.5	35.5	328	2	A44993	cytosolic repetiti	
24	94	35.3	433	2	S25194	zucotin - yeast (Sa	
25	94	35.3	703	2	T48600	kinase-like protei	
26	92.5	34.8	182	2	S61926	histone H1 homolog	
27	92.5	34.8	1128	2	T03296	R27-2 protein - Tr	
28	92	34.6	220	2	A28456	histone H1.10 - ch	
29	92	34.6	236	2	S22322	histone H1 - wheat	

30	91.5	34.4	185	2	A32137	histone H1-delta -
31	91.5	34.4	284	2	T06241	histone H1 (clone
32	91	34.2	241	2	JN0748	histone H1-II - Vo
33	90.5	34.0	581	2	E75383	conserved hypothet
34	90.5	34.0	1532	2	A26039	Iga-specific metal
35	90	33.8	231	2	S59589	histone H1 - Chlam
36	90	33.8	407	1	EDBEQ3	immediate-early pr
37	90	33.8	884	2	AI0424	translation initia
38	89.5	33.6	136	2	T14305	probable late embr
39	89.5	33.6	194	1	HSTR1	histone H1 - trout
40	89.5	33.6	356	2	A82152	toIA protein VC183
41	89.5	33.6	438	2	G87675	arylesterase-relat
42	89	33.5	1655	2	T2633	hypothetical prote
43	88.5	33.3	1482	2	T34010	hypothetical prote
44	88	33.1	288	2	T06257	histone H1 (clone
45	88	33.1	461	2	T03561	hypothetical prote

ALIGNMENTS

RESULT 1

JUV0057
 tola protein - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
 C;Accession: JUV0057; B64810
 R;Levengood, S.K.; Webster, R.E.
 J. Bacteriol. 171, 6600-6609, 1989
 A;Title: Nucleotide sequences of the tola and tolb genes and localization o
 A;Reference number: JUV0057; MUID:90078104; PMID:2687247
 A;Accession: JUV0057
 A;Molecule type: DNA
 A;Residues: 1-421 <LEV>
 A;Cross-references: UNIPROT:P19334; GB:M82832; NID:g148018; PIDN:AAA24683.1
 A;Experimental source: strain JM105
 A;Note: the authors translated the initiation codon GTG for residue 1 as Val
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503

A:Title: The complete genome sequence of *Escherichia coli* K-12. SCIENCE 277, 1433-1446, 1997

A; title: the complete genome sequence of *Escherichia coli* K-12.
A; Reference number: A64720; MUID: 97426617; PMID: 9278503
A; Accession: B64810

A: Status: nucleic acid sequence not shown: translation not shown

A:Molecule type: DN

A;Molecule type: DNA
A-Residues: 1-421 <BI.AT>

A; Residues: I-421 <BPAI>
A: Cross-references: GB: A6000177; GB: 000096; NTD: 0178

A; CROSS-REFERENCE: GB:AE000177; GB:000038; NID:G1755
A: Experimental source: strain K-12 substrain MG1655

A;Experimental source: strain K-12, substrain MG1635

U;Comment: U;Connection:

C;Genetics:

A;Gene: to1A

A;Map position: 17

A; Start codon: GTG

C;Keywords: nucleotide binding; p-loop; transmembrane

F;14-34/Domain: transmembrane #status predicted <MSS

F;78-301/Domain: helical #status predicted <HSR>

F;355-362/Region: nucleotide-binding motif A (P-loop

Query Match 44.0%; Score 117; DB 2; Length 421;

Best Local Similarity 54.7%; Pred. No. 0.0048;

Matches 35; Conservative 5; Mismatches 10; Indels

Qy 7 KEKAYAKKAEKAKKAEKAYK-----AAEAKKKEA-----KYKA

[illegible]

Db 148 KAEADAKAEAAKKAADAKKKAEAEAKAAEAQKKAEEAAALKKKA

22

53 AYE 56

—

208 BKKA 211

117 500V 002 00

PRINTED AT THE

RESULT 2

AE1317

[illegible]

Qy	1	AKKYAKKEKAYAK - KAEKAANKAEAKYKAAEAKKKKAEAKYKAAEAACAACEAAYEA	56
		: :	
Db	132	ARAAAKKAAAARKRAALAKKCAAAAAKTKAAATKAAKKPKKKKTAAKKAKKPAAKKA	188

Search completed: December 14, 2004, 05:52:13
Job time : 13.8776 secs

This Page Blank (uspto)

DR PRINTS; PR01852; SIBAPROTEIN.
DR PRINTS; PR00194; TROPOMYOSIN.
KW Complete proteome.
SQ SEQUENCE 389 AA; 42197 MW; B4032F2A2FD9B94B CRC64;

Query Match 44.0%; Score 117; DB 2; Length 389;
Best Local Similarity 55.9%; Pred. No. 0.036;
Matches 33; Conservative 7; Mismatches 15; Indels 4; Gaps 2;

Qy 1 AKKYAKKAYAK--KAEEAAKKAEAKYKAAEAKKAEAKYKAAEAKKAAEAEEA 56
||| :||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||
Db 189 AKKEAKAEAEAKAEAKV-KAEKAEAEAKV-KAEKAEAEAKAEAKAEAKAEKA 246

RESULT 5
TOLA ECOLI STANDARD; PRT; 421 AA.

AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE ToLA protein.
GN Name-toLA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KL12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levensgood S.K.; Webster R.E.;
RT "Nucleotide sequences of the tolA and tolB genes and localization of
their products, components of a multistep translocation system in
Escherichia coli.";
RT J. Bacteriol. 171:6600-6609 (1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KL12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KL12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levensgood S.K.; Beyer W.F. Jr.; Webster R.E.;
RT "Tola; a membrane protein involved in colicin uptake contains an
extended helical region.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943 (1991).
RN [5]
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R.; Gavioli M.; Benedetti H.; Prilipov A.; Lazdunski C.,
Lioudes R.;
RT "Tola central domain interacts with Escherichia coli porins.";
RL EMBO J. 15:6408-6415 (1996).

```

QY 7 KEKAYAKKAEKAAKAEKAYK-----AAEAKKKAE-----KYAEAAKAAKAA 52
Db 148 KAEADAKAEAEKAAKAAADAKKAEAEAAKAAAEAKKAEAAALKKAEAEAAAEAA 207
QY 53 AYEAE 56
Db 208 RKKA 211

RESULT 6
Q815W4
ID Q815W4 PRELIMINARY; PRT; 239 AA.
AC Q815W4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Lmo1941 protein.
GN OrderedLocusNames=lmo1941;
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Ruaniok C., Amend A.,
RA Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Duesurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krest J., Kuhn M., Kunst F., Kurapkat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordieck G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL591981; CAD00019.1; -.
DR PIR; AE1317; AE1317.
DR Listlist; LMO1941; -.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR InterPro; IPR002482; LysM.
DR Pfam; PF01476; LysM; 1.
DR SMART; SM00257; LysM; 1.
KW Complete proteome.
SQ SEQUENCE 239 AA; 25836 MW; 72E59D576E0D7832 CRC64;

Query Match 43.8%; Score 116.5; DB 2; Length 239;
Best Local Similarity 55.4%; Pred. No. 0.026;
Matches 31; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAEKAAKAEKAAKAEKAAKAEKAAKAE 55
Db 116 AEKAEKAAAEKAAAEKADKKQBEDAVKAAKAEKAEKAAAEKAAAEKAAAE 171

RESULT 7
Q6N503
ID Q6N503 PRELIMINARY; PRT; 105 AA.
AC Q6N503;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RP3180;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;

Query Match 43.8%; Score 116.5; DB 2; Length 239;
Best Local Similarity 55.4%; Pred. No. 0.026;
Matches 31; Conservative 9; Mismatches 15; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAEKAAKAEKAAKAEKAAKAEKAAKAE 55
Db 116 AEKAEKAAAEKAAAEKADKKQBEDAVKAAKAEKAEKAAAEKAAAEKAAAE 171

RESULT 8
CAE28621
ID CAE28621 PRELIMINARY; PRT; 105 AA.
AC CAE28621;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN RPA3180.
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CAE28621.1; -.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 105 AA; 11042 MW; CEDB59B3D937E980 CRC64;

Query Match 43.4%; Score 115.5; DB 2; Length 105;
Best Local Similarity 58.6%; Pred. No. 0.016;
Matches 34; Conservative 4; Mismatches 17; Indels 3; Gaps 2;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAEKAAKAEKAAKAEKAAKAEKAAKAE 56
Db 33 AKKVAKKVAGAKKAAKATKKAGKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 89

RESULT 9
Q7RC08
ID Q7RC08 PRELIMINARY; PRT; 1701 AA.
AC Q7RC08;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Erythrocyte binding protein.
GN Name=PY05977;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]

```

```

RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CAE28621.1; -.
DR InterPro; IPR005819; Histone_H5.
DR PRINTS; PR00624; HISTONEH5.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 105 AA; 11042 MW; CEDB59B3D937E980 CRC64;

Query Match 43.4%; Score 115.5; DB 2; Length 105;
Best Local Similarity 58.6%; Pred. No. 0.016;
Matches 34; Conservative 4; Mismatches 17; Indels 3; Gaps 2;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAEKAAKAEKAAKAEKAAKAEKAAKAE 56
Db 33 AKKVAKKVAGAKKAAKATKKAGKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 89

```

[illegible]

```
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL: AE016757; AN79291.1; -.
DR HSSP: P19934; ITOL.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 421 AA; 43184 MW; DB296626F056D385 CRC64;

Query Match 43.2%; Score 115; DB 2; Length 421;
Best Local Similarity 56.2%; Pred. No. 0.056;
Matches 36; Conservative 6; Mismatches 8; Indels 14; Gaps 3;

QY 7 KEKAYAKAEKAA-----KKAEEKAYKAA-EAKKKAEAE-----KYKAEAAKAAKAE 52
Db 148 KAEADAKAEAEAKKAAADAKKAEAEAEAKAAVAEQKAEAEAAALKKAEAEAEAAEA 207
QY 53 AYEAE 56
Db 208 RKKA 211

RESULT 13
Q7W477 PRELIMINARY; PRT; 373 AA.
AC Q7W477
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proline-rich inner membrane protein.
GN OrderedLocuNames=BPP3791;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40 (2003).
DR EMBL: BX640434; CAE39074.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
DR GO: GO:0008565; F:protein transport activity; IEA.
DR GO: GO:0015031; P:protein transport; IEA.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
DR TIGRFAMs: TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 373 AA; 40092 MW; A3648994DAE46E328 CRC64;

Query Match 43.0%; Score 114.5; DB 2; Length 373;
Best Local Similarity 55.4%; Pred. No. 0.055;
Matches 31; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 1 AKKYAKKEKAYAKA-EKAAKAEKAYKAEAKKAEAKYKAEAKAAKAAKAEAYE 55
Db 182 AEKAAEKQAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 237

RESULT 14
Q937K4 PRELIMINARY; PRT; 395 AA.
AC Q937K4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tola protein.
GN Name=tola;
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=3937;
RA Ray M.C., Vianney A., Cotte-pattat N., Lazzaroni J.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ297885; CAC82708.1; -.
DR HSSP: P19934; ITOL.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
SQ SEQUENCE 395 AA; 41601 MW; 3C0C1DC12E181013 CRC64;

Query Match 43.0%; Score 114.5; DB 2; Length 395;
Best Local Similarity 54.2%; Pred. No. 0.058;
Matches 32; Conservative 11; Mismatches 13; Indels 3; Gaps 2;

QY 1 AKKYAKKE-KAYAKAEKAAKKA--EAKYKAAEAKKKAEAKYKAEAKAAKAAKAEAYE 56
Db 182 AKKAAEEAEKTKAAAAAEAKKAAEEAEKAAKAAADAKQKAAEEAEKAAKAAEAKKAAADA 240

RESULT 15
Q8ZQT6 PRELIMINARY; PRT; 407 AA.
AC Q8ZQT6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tol protein, membrane spanning protein.
GN Name=tola; OrderedLocuNames=STM0747;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL: AE008730; AAL19691.1; -.
DR HSSP: P19934; ITOL.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 407 AA; 41865 MW; 2534352116602D75 CRC64;

Query Match 42.7%; Score 113.5; DB 2; Length 407;
Best Local Similarity 49.3%; Pred. No. 0.072;
Matches 37; Conservative 10; Mismatches 9; Indels 19; Gaps 4;

QY 1 AKKYAKKEKAYAKKAEKAA---KKAEEKAYK-AAEAKKKAEAE-----KYKAEAE 44
Db 123 AKLAQQQQQAEAEKAAADAKKAAEAEKAAKAAADAKKKAAEAEKAAADAKKAAEAE 182
QY 45 AKAAA---KEAEAYE 56
```


Db ||||| |:| ||
 183 AKAAADAKKKAEEA 197

Search completed: December 14, 2004, 06:10:16
Job time : 63.8523 secs

This Page Blank (uspto)

Eur. J. Biochem. 225, 1089-1095, 1994
A:Title: Tandemly arranged repeats of a novel highly charged 16-amino-acid motif represses the growth of the extremely elongated spermatozoa of Drosophila hydei.
A:Reference number: S51364; MUID:95045538; PMID:7957199
A:Accession: S51364
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1390 <NEE>
A:Cross-references: UNIPROT:Q08696; EMBL:X73481
R:Neesen, J.; Heinlein, U.A.O.; Buenemann, H. submitted to the EMBL Data Library, June 1993
A:Reference number: S34153
A:Accession: S34154
A:Molecule type: DNA
A:Residues: 1-163, 'E', 164-236, 'Q', 237-254, 257-320, 'E', 321-1390 <NEW>
A:Cross-references: EMBL:X73481; NID:g313201; PID:g313202
C:Genetics:
A:Gene: mst101(2)
A:Cross-references: FlyBase:FBgn0011816

Query Match 38.8%; Score 121.5; DB 2; Length 1390;
Best Local Similarity 59.7%; Pred. No. 0.022;
Matches 37; Conservative 3; Mismatches 19; Indels 3; Gaps 3;

QY 2 KKYAKKRYAKAKAEAKKAAKAAKAEKKYAKAAKAEKKYAAAEAKYAAAEAKAAAKE 61
DB 655 KKLAKKETAEKKCE-KAAKKRKEAEK-KKCAEAAKKEAEAKKKCEAAK-KEKE 711
QY 62 AA 63
DB 712 AA 713

RESULT 7
E81525
TolA protein PA0971 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: E81525
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Brlaganza, S.; Kuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E81525
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <STO>
A:Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG0436
A:Experimental source: strain PA01
C:Genetics:
A:Gene: tolA; PA0971

Query Match 38.2%; Score 119.5; DB 2; Length 347;
Best Local Similarity 51.5%; Pred. No. 0.011;
Matches 34; Conservative 7; Mismatches 22; Indels 3; Gaps 1;

QY 1 AKKYAK---KEKAYAKAKAEAKKAAKAAKAEKKYAKAAKAEKKYAAAEAKYKAEAKA 57
DB 155 AKKAEDEAKKAAEDAKKAAEDAKKAAEBAKKAEEAKKAAAEAAKKAEEAKKGAAGAAAAA 214
QY 58 AAKEAA 63
DB 215 AAKKAA 220

RESULT 8
S59589
histone H1 - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59589; S62122

R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Cornelius, T.; Schmitt, R. Curr. Genet. 28, 333-345, 1995
A:Title: The organization, structure and regulatory elements of Chlamydomonas histone gene
A:Reference number: S59581; MUID:96120862; PMID:8590479
A:Accession: S59589
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-231 <FAB>
A:Cross-references: UNIPROT:Q39576; EMBL:U16726
A:Note: the authors did not translate the codon for residue 1
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Cornelius, T.; Schmitt, R. submitted to the EMBL Data Library, October 1994
A:Description: The organization, structure and controlling elements of Chlamydomonas histone gene
A:Reference number: S62122
A:Accession: S62122
A:Molecule type: DNA
A:Residues: 1-173, 'P', 174-231 <FAW>
A:Cross-references: EMBL:U16726; NID:g571479; PIDN:AAA98452.1; PID:g571480
C:Genetics:
A:Introns: 62/3; 101/3
C:Superfamily: histone H1
C:Keywords: chromosomal protein; DNA binding; nucleosome

Query Match 37.7%; Score 118; DB 2; Length 231;
Best Local Similarity 55.6%; Pred. No. 0.01;
Matches 35; Conservative 5; Mismatches 21; Indels 2; Gaps 2;

QY 3 KYAKKEKAYAKAEAKKAAKAEAK-KYAKAAKAEKKYAA-AEAKYKAEAAKAAAK 60
DB 159 KAEKPKAAKPAKTKTKAAAKPKAEKPKAAKPKAEKPKAAKPKAEKPKAAKPAK 218
QY 61 EAA 63
DB 219 KSA 221

RESULT 9
G64064
outer membrane integrity protein tolA - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: G64064; JC5212
R:Felschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: G64064
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <TIGR>
A:Cross-references: UNIPROT:P44678; GB:U32722; GB:L42023; NID:gl573348; PIDN:AAC22041.1;
R:Sen, K.; Sikkema, D.J.; Murphy, T.F. Gene 178, 75-81, 1996
A:Title: Isolation and characterization of the Haemophilus influenzae tolQ, tolR, tolA and tolB genes.
A:Reference number: JC5212; MUID:97080550; PMID:8921895
A:Accession: JC5212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'V', 2-47, 'A', 49-141, 'R', 143-164, 'P', 166-189, 'R', 191-202, 'A', 204-226, 'A', 228-231
A:Cross-references: GB:U32470
A:Experimental source: strain 1479
A:Note: the authors translated the codon CGT for residue 190 as Ala
C:Genetics:
A:Gene: tolA
A:Start codon: GTG

Query Match 37.5%; Score 117.5; DB 2; Length 372;
Best Local Similarity 46.6%; Pred. No. 0.016;
Matches 34; Conservative 10; Mismatches 22; Indels 7; Gaps 1;

Mol. Cell. Biol. 7, 478-485, 1987

A>Title: Isolation, characterization, and expression of the gene encoding the late histone H1

A;Reference number: A26721; MUID:87172742; PMID:3031476

A;Accession: A26721

A:Molecule type: DNA

A;Residues: 1-217 <KNO>

A;Cross-references: UNIPROT:P07796; GB:M16033; NID:g161517; PIDN:AAA30059.1; PID:g161518

C:Superfamily: histone H1

C;Keywords: chromosomal protein; DNA binding; nucleosome; nucleus

F;2-217/Product: histone H1-gamma, embryonic #status predicted <MAT>

Query Match 35.9%; Score 112.5; DB 2; Length 217;
Best Local Similarity 49.4%; Pred. No. 0.026; Indels 15; Gaps 3;
Matches 38; Conservative

Qy 1 AKKYAKKE-----AYAKAKAEAAK-----AAFKAKAEAAKTKVAKAAK---AEKKYEA 45
||| ||| ||| ||| ||| ||| ||| ||| : : :
Db 139 AKKATKTKTKVKVPAAKKAACKPAAKPPAAKPAKKAAPAAKPAKPAKPAKPAKPAKPAKPAK 198
||| ||| ||| ||| ||| ||| ||| ||| : : :
Qy 46 AEAKYKAEAAKAAAKEA 62
||| : ||| |||
Db 199 KPAAKAAKPAAKAAKPA 215

RESULT 15

AE1689
hypochemical protein homolog lin2055 [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1689
R:Glaser, P.; Frangaul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefef, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Makarewicz, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, O.; Title: Comparative Genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1689
A:Status: preliminary
A:Molecule type: DNA
A:Molecule type: DNA
A:Residues: 1-243 <LA>
A:Cross-references: UNIPROT:Q92A67; GB:ALU592022; PIDN:CAC97285.1; PID:gl6414556; GSPDB:G1689
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2055

```
Query Match      35.8%; Score 112; DB 2; Length 243;
Best Local Similarity 47.1%; Pred. No. 0.031;
Matches 32; Conservative 11; Mismatches 23; Indels 2; Gaps 1;

Qy   1 AKTYAKGKAVAKAAEAKAQAQAABAKKYAKAAAEKEYAAAAE--AKYKAEEAAKAA 58
      : : | | | : | | | | | : : | | | : : | | | : : | | |
Db 110 AEEXAAAEKAAEBKAAEAKEAAADKKSQEDESAAKAAAKKQGEAAEEKAAAADKAA 169
      : : | | | : | | | | | : : | | | : : | | | : : | | |

Qy   59 AKEAAAYEA 66
      : : | | |
Db 170 KEKAAAEA 177
```

Search completed: December 14, 2004, 05:52:13
Job time : 15.1772 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	132.5	42.3	394	2	Q7AGI8	Q7ag18 escherichia
2	132.5	42.3	394	2	Q8X965	Q8x965 escherichia
3	131	41.9	421	1	TOLA_ECOLI	P19934 escherichia
4	128	40.9	421	2	Q8FUT1	Q8fjt1 escherichia
5	127.5	40.7	413	2	Q7C2Q4	Q7c2q4 shigella fl
6	127.5	40.7	413	2	Q83SA1	Q83sa1 shigella fl
7	126.5	40.4	389	2	Q9CM70	Q9cm70 pasteurella
8	124.5	39.8	407	2	Q8ZOT6	Q8zot6 salmonella
9	123	39.3	248	1	HL_PARAN	P02256 parechinus
10	122.5	39.1	372	2	Q9WXX1	P02256 pseudomonas
11	122.5	39.1	372	2	Q88NI6	Q88ni6 pseudomonas
12	122.5	39.1	376	2	Q828C1	Q828c1 salmonella
13	122	39.0	713	2	Q6CDX0	Q6cdx0 yarrowia li
14	121.5	38.8	401	2	Q74W64	Q74w64 yersinia pe
15	121.5	38.8	401	2	AAS61283	Aas61283 yersinia
16	121.5	38.8	1391	1	MST2_DROHY	Mst2 drohy
17	120	38.3	395	2	Q937K4	Q937k4 erwinia chr
18	120	38.3	395	2	Q6D7F3	Q6d7f3 erwinia car
19	120	38.3	1701	2	Q7RC08	Q7rc08 plasmodium
20	119.5	38.2	347	1	TOLA_PSEAE	P50600 pseudomonas
21	118	37.7	177	2	Q6SG84	Q6sg84 uncultured
22	118	37.7	177	2	AAR37978	Aar37978 uncultured
23	117.5	37.5	372	1	TOLA_HAEIN	P44678 haemophilus
24	117	37.4	356	2	Q7NGT7	Q7ngt7 photorhabdu
25	117	37.4	388	2	Q8ZGZ2	Q8zgz2 yersinia pe
26	117	37.4	393	2	Q8CZ28	Q8cz28 yersinia pe
27	116	37.1	373	2	Q7W477	Q7w477 bordetella
28	115.5	36.9	206	2	Q7MA09	Q7ma09 parechinus
29	115.5	36.9	379	2	Q7WFn5	Q7wfn5 bordetella
30	115	36.7	344	1	MST1_DROHY	Q88695 droso
31	115	36.7	356	2	O87Y39	Q87y39 pseudomonas

[illegible]

```

SQ SEQUENCE 413 AA; 42355 MW; 93E10F2C5DB60DE8 CRC64;
  Query Match 40.7%; Score 127.5; DB 2; Length 413;
  Best Local Similarity 56.7%; Pred. No. 0.013;
  Matches 38; Conservative 6; Mismatches 20; Indels 3; Gaps 2;
QY 1 AKKYAKKEKAYAKAKAEAKKAAKAEAKKAAKAEAKK--FYAAAEAKYKAEAAKAAA 59
Db 139 AKAADAKAAEEAAKKAADAKKKAEAAK--AAAEAKKAEVAAALKKAEAAEAAA 196
QY 60 KEAAYEA 66
Db 197 AEARKKA 203

RESULT 7
Q9CM70 PRELIMINARY; PRT; 389 AA.
AC Q9CM70;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ToLA.
GN Name:toLa; OrderedLocusNames=PM0968;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B. J., Zhang Q., Li L. L., Paustian M. L., Whittam T. S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465 (2001).
DR EMBL; AS006136; AAK03052.1; -.
DR HSSP; P01096; IGMJ.
DR InterPro; IPR009148; Siba.
DR InterPro; IPR010528; ToLa.
DR Pfam; PF06519; ToLa; 1.
DR PRINTS; PR01852; SIBAPROTEIN.
DR PRINTS; PR00194; TROPOMYOSIN.
KW Complete proteome.
SQ SEQUENCE 389 AA; 42197 MW; B4032F2A2FD9E94B CRC64;
  Query Match 40.4%; Score 126.5; DB 2; Length 389;
  Best Local Similarity 52.0%; Pred. No. 0.014;
  Matches 39; Conservative 9; Mismatches 12; Indels 15; Gaps 3;
QY 1 AKKYAKKEKAYAKAK-----KAEAKA--AKKAKAEAKKYAKAEAKKEKYAAAEAKYK 51
Db 189 AKKAEKAEAEAKAEAKVEKAEKAEAKVKEAKAEAEAKAEAKAEAK-----AEAKAK 242
QY 52 AEAKAAAEAEAYEA 66
Db 243 AEKAKADAEAAQRKA 257

RESULT 8
Q8ZQT6 PRELIMINARY; PRT; 407 AA.
AC Q8ZQT6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tol protein, membrane spanning protein.
GN Name:toLa; OrderedLocusNames=STM0747;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K. B., Spieth J., Clifton S. W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R. K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856 (2001).
DR EMBL; AE008730; AAL19691.1; -.
DR HSSP; P19934; 1TOL.
DR InterPro; IPR010528; ToLa.
DR Pfam; PF06519; ToLa; 1.
KW Complete proteome.
SQ SEQUENCE 407 AA; 41865 MW; 2534352116602D75 CRC64;
  Query Match 39.8%; Score 124.5; DB 2; Length 407;
  Best Local Similarity 47.8%; Pred. No. 0.022;
  Matches 43; Conservative 6; Mismatches 10; Indels 31; Gaps 4;
QY 1 AKKYAKKEKAYAKA---KAEAKAAK-----KAKAEAKKYAKAEAKKEY 43
Db 189 AKKAEAEAAKAAAEAKKAEAEAAKAAAEAKKADAEAAKAAAEAKKADAAAAAK---- 244
QY 44 AAAEAKYKAEAA-----KAAAEKAA 63
Db 245 AAAEAKKADAAAKAAADAKKKAAAEKAA 274

RESULT 9
H1_PARAN STANDARD; PRT; 248 AA.
AC P02256;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histone H1, gonadal.
OS Parechinus angulosus (Angulate sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoidae; Echinidae; Parechinus.
OX NCBI_TaxID=7658;
RN [1]
RP SEQUENCE OF 1-84.
RX MEDLINE=80156831; PubMed=6767609;
RA Strickland W. N., Strickland M., de Groot P. C., von Holt C.,
RA Wittmann-Liebold B.;
RT "The primary structure of histone H1 from sperm of the sea urchin
RT Parechinus angulosus. 1. Chemical and enzymatic fragmentation of the
RT protein and the sequence of amino acids in the four N-terminal
RT cyanogen bromide peptides.";
RN [2]
RP SEQUENCE OF 80-248.
RX MEDLINE=80156832; PubMed=7363905;
RA Strickland W. N., Strickland M., Brandt W. F., von Holt C., Lehmann A.,
RA Wittmann-Liebold B.;
RT "The primary structure of histone H1 from sperm of the sea urchin
RT Parechinus angulosus. 2. Sequence of the C-terminal CNBr peptide and
RT the entire primary structure.";
RL Eur. J. Biochem. 104:567-578 (1980).
CC -|- FUNCTION: Histones H1 are necessary for the condensation of
CC nucleosome chains into higher order structures.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- TISSUE SPECIFICITY: Sperm.
CC -|- SIMILARITY: Belongs to the histone H1/H5 family.
DR PIR; A91090; HSURIP.
DR HSSP; P02259; 1HST.
DR InterPro; IPR005818; Histone_H1/H5.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR003216; Linkerhist_N.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.

```

DR Pfam; PF00538; Linker_histone; 1.
DR PRINTS; PR00624; HISTONEH5.
DR ProDom; PD000373; Linkerhist_N; 1.
DR SMART; SM00526; H15; 1.
KW Chromosomal protein; Direct protein sequencing; DNA-binding;
KW Multigene family; Nuclear protein; Sperm.
FT VARIANT 144 144
SQ SEQUENCE 248 AA; 26387 MW; 1B25BFJF136541947 CRC64;

Query Match 39.3%; Score 123; DB 1; Length 248;
Best Local Similarity 56.1%; Pred. No. 0.019;
Matches 37; Conservative 7; Mismatches 20; Indels 2; Gaps 2;

OY 3 KYAKKEKAVAKAEAAKAAGA-KAEAKKYAKAAGKEKYAA-AEAKYKAEAAKAAK 60
||| ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
120 KRAKTSAAKAKAKAAAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKPKK 179

Db 61 EAAYEA 66
:||| : ||| :
180 KAAXKA 185

RESULT 10
Q9WWX1 PRELIMINARY; PRT; 372 AA.

ID	Q9WWX1	PRELIMINARY;	PRT;	372 AA.
AC	Q9WWX1			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Tola protein.			
GN	Name=tola;			
OS	Pseudomonas putida.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Pseudomonadaceae; Pseudomonas.			
OX	NCBI_TaxID=303;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=mt-2;			
RX	MEDLINE=96198174; PubMed=8626299;			
RA	Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;			
RT	"The Pseudomonas putida peptidoglycan-associated outer membrane			
RT	lipoprotein (PAL) is involved in maintenance of the integrity of the			
RT	cell envelope.";			
RT	J. Bacteriol. 178:1699-1706(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=mt-2;			
RX	MEDLINE=96422022; PubMed=8824639;			
RA	Rodriguez-Herva J.J., Ramos J.;			
RT	"Characterization of an OprL null mutant of Pseudomonas putida.";			
RT	J. Bacteriol. 178:5836-5840(1996).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=mt-2;			
RA	Ramos-Gonzalez I.;			
RL	Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.			
RL	EMBL; X74218; CAB50780.1; -.			
DR	HSSP; P50600; 1LR0.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0000786; C:nucleosome; IEA.			
DR	GO; GO:0005634; C:nucleus; IEA.			
DR	GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . . ; IEA.			
DR	GO; GO:0003677; F:DNA binding; IEA.			
DR	GO; GO:0008565; F:protein transporter activity; IEA.			
DR	GO; GO:0006334; P:nucleosome assembly; IEA.			
DR	GO; GO:0015031; P:protein transport; IEA.			
DR	InterPro; IPRO05819; Histone_H5.			
DR	InterPro; IPRO10528; TOLA.			
DR	Pfam; PF06519; TOLA; 1.			
DR	PRINTS; PR00624; HISTONEH5.			
DR	TIGRFAMe; TIGR01352; tonB_Cterm; 1.			
KW	Complete proteome.			
SQ	SEQUENCE 372 AA; 40133 MW; 87F49785EC3C0BC CRC64;			

Query Match 39.1%; Score 122.5; DB 2; Length 372;
Best Local Similarity 51.5%; Pred. No. 0.029;
Matches 35; Conservative 10; Mismatches 18; Indels 5; Gaps 2;

OY 1 AKKYAKEKAVAKAEAAKAAGA-EAAKKYAKAAGKEKYAA-EAAKAA 58
||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
148 AKAAEKQAADIANKKAEDEAKKAEAEAK---KAAAEAAKKAADAKKAAEEAKKA 204

Db 59 AERAYEA 66
:||| : ||| :
205 AEDAKKA 212

RESULT 11
Q88NI6 PRELIMINARY; PRT; 372 AA.

ID	Q88NI6	PRELIMINARY;	PRT;	372 AA.
AC	Q88NI6			
DT	01-JUN-2003 (TrEMBLrel. 24, Created)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Biopolymer transport protein Tola.			
GN	Name=tola; OrderedLocNames=PP1221;			
OS	Pseudomonas putida (strain KT2440).			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;			
OC	Pseudomonadaceae; Pseudomonas.			
OX	NCBI_TaxID=160488;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=22423060; PubMed=12534463;			
RA	Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,			
RA	Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,			
RA	Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,			
RA	Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,			
RA	Hance I., Chris Lee P., Holtzapfel E.K., Scanlan D., Tran K.,			
RA	Moazed A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,			
RA	Weitzer H., Lauber J., Stjepandic D., Honeisel J., Straetz M., Heim S.,			
RA	Kiewitz C., Eisen J.A., Timmis K.N., Duisterhoef A., Tuemmler B.,			
RA	Fraser C.M.;			
RT	"Complete genome sequence and comparative analysis of the			
RT	metabolically versatile Pseudomonas putida KT2440.";			
RL	Environ. Microbiol. 4:799-808(2002).			
DR	EMBL; AE016778; AAN66845.1; -.			
DR	HSSP; P50600; 1LR0.			
DR	TIGR; PP1221; -.			
DR	GO; GO:001			

```

Db 148 AKKAAKQADIAKKAADEAKKABEEAK---KAAAEBAKKAEDAKKAAEEAKKKA 204
QY 59 AKEAAVEA 66
Db 205 AEDAKKKA 212

RESULT 12
Q828C1 PRELIMINARY; PRT; 376 AA.
AC Q828C1; Q7C8P3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tola protein.
GN Name:tola; OrderedLocusNames-STV0793, t2129;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TV2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyonianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AL627268; CAD05209.1; -.
DR EMBL; AE016841; ARO69743.1; -.
DR HSSP; P19934; ITOL.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 376 AA; 38804 MW; EC21P2C4767A8A42 CRC64;

Query Match 39.1%; Score 122.5; DB 2; Length 376;
Best Local Similarity 51.3%; Pred. No. 0.029;
Matches 40; Conservative 4; Mismatches 13; Indels 21; Gaps 3;

QY 1 AKKYAKKEKAYAKA---KKAEEAAK-----KAKAAKKYAKAAKAEKKEY 43
Db 173 AKKAAEAAKAAAEAKKGAEEAAKAAADAKKKADAAEAAKKAADAAAK---- 228

QY 44 AAAAEAKYKAAEAAKAAAKE 61
Db 229 AAADAKKAAAEKAAAE 246

RESULT 13
Q6CDX0 PRELIMINARY; PRT; 713 AA.
AC Q6CDX0;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to sp|P38678 Neurospora crassa Glucan synthase-1.

```

```

GN ORFNames=YALI0B20570G;
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG GENOLEVURES;
RA Dujon B., Sherman D., Fischer G., Durrans P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattolico L., Confanieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382128; CAG83395.1; -.
SQ SEQUENCE 713 AA; 77807 MW; 5DD84D2C34CF55AC CRC64;

Query Match 39.0%; Score 122; DB 2; Length 713;
Best Local Similarity 50.7%; Pred. No. 0.056;
Matches 35; Conservative 8; Mismatches 18; Indels 8; Gaps 2;

QY 2 KKVAKEKAYAKAKKAE-----AKAAKAAKAAK--KYAKAAKAEKVEAAAEKVKAE 53
Db 571 EKEEKAKEAAKAAKAEELKEEVAKAAKAESEQATAAAEAAKAAKAAEADASKKVE 630

QY 54 AAKAAAKEA 62
Db 631 AEKAAAEES 639

RESULT 14
Q74W64 PRELIMINARY; PRT; 401 AA.
AC Q74W64;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Tola colicin import membrane protein.
GN Name:tola; OrderedLocusNames=YP1033;
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=91001 / Biovar Mediaevalis;
RA Song Y., Tong Z., Wang L., Han Y., Zhang J., Pei D., Wang J., Zhou D.,
RA Han Y., Pang X., Zhai J., Chen F., Qin H., Wang J., Li S., Guo Z.,
RA Ye C., Du Z., Lin W., Wang J., Yu J., Yang H., Wang J., Huang P.,
RA Yang R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017130; AAS61283.1; -.
DR InterPro; IPR010528; Tola.
DR Pfam; PF06519; Tola; 1.
SQ SEQUENCE 401 AA; 41868 MW; 43D682DE91CF1301 CRC64;

Query Match 38.8%; Score 121.5; DB 2; Length 401;

```


This Page Blank (uspto)

[illegible]


```
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59321
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59321

Query Match      40.8%; Score 108.5; DB 15; Length 323;
Best Local Similarity 55.7%; Pred.No. 0.01; Indels 7; Gaps 3;
Matches 34; Conservative 7; Mismatches 13; Indels 7; Gaps 3;

QY 1 AKKYAKKAYAKKAAKAAKAAKAYK-AAEAKKAAEA----KYKAAEAAKAAKAAKAAEAAAYE 55
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 65 AKKQA--EAAAKAAAEAKKQAAEAAKAAKAAEAAKAAEAAKAAKAAKAAKAAKAAKAAKAAKQ 122
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 56 A 56
    |
Db 123 A 123
```

Search completed: December 14, 2004, 06:59:41
Job time : 139.937 secs

This Page Blank (uspto)


```

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match 41.7%; Score 111; DB 15; Length 428;
Best Local Similarity 53.1%; Pred. No. 0.008;
Matches 34; Conservative 11; Mismatches 11; Indels 8; Gaps 3

QY 1 AKKVA--KKEKAYAKKAERKAR---KKAEAKYK--AAEAKKCAEAKYKAEAAKAAKEA 52
Db 167 AKKAADAQKKAEAEAKKAAADAQKKAEAEAKKAAADAQKKAEAEAKKAAQAQAEKKA 226

QY 53 AYEAE 56
Db 227 AAEA 230

RESULT 14
US-10-282-122A-75772
; Sequence 75772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

```

```

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72645
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Salmonella paratyphi A
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (78)..(78)
; OTHER INFORMATION: X=any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (303)..(303)
; OTHER INFORMATION: X=any amino acid
; US-10-282-122A-72645

Query Match 42.3%; Score 112.5; DB 15; Length 387;
Best Local Similarity 53.5%; Pred. No. 0.0052;
Matches 38; Conservative 5; Mismatches 11; Indels 17; Gaps 4;

QY 1 AKKYAKKAYAKKAAKAA-----KKAEEKAYK-AAEAKKKAA-----KYKAAKAA 48
Db 141 AKKGAEPAA--AKAAADAKKAAEAATKAAADAKKKAAEAATAAADAADAKKKAAEAATAA 198
QY 49 A---KEAAVEA 56
Db 199 AEAKKKAAEA 209

RESULT 13
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

```

[illegible]

Result No.	Score	Query #			DB	ID	Description
		Match	Length				
1	266	100.0	56	9	US-09-816-989A-3	Sequence 3, Appli	
2	245.5	92.3	77	9	US-09-816-989A-5	Sequence 5, Appli	
3	233	87.6	86	9	US-09-816-989A-6	Sequence 6, Appli	
4	198	74.4	66	9	US-09-816-989A-4	Sequence 4, Appli	
5	180.5	67.9	109	9	US-09-816-989A-7	Sequence 7, Appli	
6	134.5	50.6	45	9	US-09-816-989A-2	Sequence 2, Appli	
7	117	44.0	372	15	US-10-282-122A-68109	Sequence 68109, A	
8	117	44.0	389	15	US-10-282-122A-67145	Sequence 67145, A	
9	117	44.0	421	15	US-10-282-122A-56483	Sequence 56483, A	
10	116.5	43.8	239	15	US-10-282-122A-60543	Sequence 60543, A	
11	113.5	42.7	407	15	US-10-282-122A-75047	Sequence 75047, A	
12	112.5	42.3	387	15	US-10-282-122A-72645	Sequence 72645, A	
13	111	41.7	428	15	US-10-282-122A-55748	Sequence 55748, A	

Search completed: December 14, 2004, 06:01:12
Job time : 62.9072 secs

```
PS Claim 25; SEQ ID NO 75047; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 407 AA;

Query Match 42.7%; Score 113.5; DB 6; Length 407;
Best Local Similarity 49.3%; Pred. No. 0.0014;
Matches 37; Conservative 10; Mismatches 9; Indels 19; Gaps 4;

QY 1 AKKYAKKEKAYAKKAEKAA-----KKEAKYK-AAEAKKAEKAE-----KYKAEK 44
Db 123 AAKLAQQQQQAEAKYAAADAKKAEKAEKAAADAKKAEKAEKAAADAKKAEKAEK 182
QY 45 AKAAA---KEAAYEA 56
Db 183 AKAAADAKKAEKAE 197

RESULT 13
ABU44721
ID ABU44721 standard; protein; 387 AA.
XX
XX ABU44721;
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
XX
XX 08-FEB-2002; 2002US-00072851.
XX

PR Claim 25; SEQ ID NO 72645; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 387 AA;

Query Match 42.3%; Score 112.5; DB 6; Length 387;
Best Local Similarity 53.5%; Pred. No. 0.0016;
Matches 38; Conservative 5; Mismatches 11; Indels 17; Gaps 4;

QY 1 AKKYAKKEKAYAKKAEKAAKAEKAYK-AAEAKKAEKAE-----KYKAEKAAK 48
Db 141 AKKKAEPKAE--AKAAADAKKAEKAEKAAADAKKAEKAEKAAADAKKAEKAEK 198
QY 49 A---KEAAYEA 56
Db 199 AEAKKAEKAE 209

RESULT 14
AAR06446
ID AAR06446 standard; protein; 106 AA.
XX
XX AAR06446;
XX
XX 25-MAR-2003 (revised)
XX
XX 03-JAN-1991 (first entry)
XX
XX Recombinant copolymer 1-19, myelin basic protein analogue.
XX
```



```

XX OS Unidentified.
XX KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
XX KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; antipruritic; dermatological;
XX KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
XX KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
XX PD Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
XX KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
XX KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
XX KW pemphigus vulgaris; systemic lupus erythematosus.
XX OS Unidentified.
XX KW WO200018794-A1.
XX PN WO200018794-A1.
XX PD 06-APR-2000.
XX PF 24-SEP-1999; 99WO-US022402.
XX PR 25-SEP-1998; 98US-0101693P.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (TEVA-) TEVA PHARM USA INC.
XX PI Gad A, Lis D;
XX WPI; 2000-317499/27.
XX Copolymer 1 related polypeptides used as molecular weight markers for
XX glatiramer acetate and for treatment and prevention of immune diseases.
XX Claim 10; Page 14; 72pp; English.
XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
XX weight TV-marker polypeptides from the present invention. The present
XX invention describes polypeptides (I) for determining the molecular weight
XX of a copolymer (CP), which has an identified molecular weight and an
XX amino acid composition corresponding to the copolymer. The polypeptides
XX of the invention are used as molecular weight markers for glatiramer
XX acetate related tetrapolymers. The polypeptides may also be used for
XX treating and preventing immune diseases in a mammal. Autoimmune diseases
XX which may be treated include either cell-mediated or antibody-mediated
XX diseases. Such diseases include arthritic conditions, demyelinating
XX diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
XX arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
XX oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
XX disease, chronic immune thrombocytopaenia purpura, colitis, contact
XX sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
XX syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
XX psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
XX diseases which can be treated include host-versus-graft disease,
XX graft-versus-host disease, and delayed-type hypersensitivity. The
XX polypeptides of the invention have defined molecular weights and physical
XX properties which are analogous to glatiramer acetate molecules, which
XX makes them ideal for use as molecular weight markers
XX Sequence 109 AA;
XX
XX Query Match 67.9%; Score 180.5; DB 3; Length 109;
XX Best Local Similarity 45.9%; Pred. No. 5.8e-11;
XX Matches 50; Conservative 3; Mismatches 3; Indels 53; Gaps 3;
XX
XX QY 1 AKKYAKK-EKAYAKKA-----EKAAYKAEKAYKAEAKKAEAK----- 39
XX Db 1 AKKYAKKAEKAYAKKAKEKAYAKKAEKAYAKKAEKAEAKKAEAKKAEAKKEA 60
XX
XX QY 40 -----YKAEAAKAAKAEAAKAEAAKAEAAKAEAAKAEAAKAEAAKAEAA 56
XX Db 61 YKAEAKKYAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 109
XX
XX RESULT 6
XX AAY82572
XX ID AAY82572 standard; peptide; 45 AA.
XX AC AAY82572;
XX XX
XX XX 28-JUL-2000 (first entry)
XX DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:2.
XX KW Copolymer; molecular weight marker; TV-marker; immune disease;
XX KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 61.9072 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-3

Perfect score: 266

Sequence: 1 AKKYAKKEKAVAKAEAKA.....EAKYKAEAKAAKEAYEA 56

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23sep04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	266	100.0	56	3	AAY82573	Aay82573 Copolymer
2	245.5	92.3	77	3	AAY82575	Aay82575 Copolymer
3	233	87.6	86	3	AAY82576	Aay82576 Copolymer
4	198	74.4	66	3	AAY82574	Aay82574 Copolymer
5	180.5	67.9	109	3	AAY82577	Aay82577 Copolymer
6	134.5	50.6	45	3	AAY82572	Aay82572 Copolymer
7	117	44.0	372	6	ABU40185	Abu40185 Protein e
8	117	44.0	389	6	ABU39221	Abu39221 Protein e
9	117	44.0	421	6	ABU28559	Abu28559 Protein e
10	116.5	43.8	239	5	ABB49123	Abb49123 Listeria
11	116.5	43.8	239	6	ABU32619	Abu32619 Protein e
12	113.5	42.7	407	6	ABU47123	Abu47123 Protein e
13	112.5	42.3	387	6	ABU44721	Abu44721 Protein e
14	111	41.7	106	2	AAR06446	Aar06446 Recombina
15	111	41.7	428	6	ABU27824	Abu27824 Protein e
16	110.5	41.5	376	6	ABU47848	Abu47848 Protein e
17	108.5	40.8	323	6	ABU31397	Abu31397 Protein e
18	108.5	40.8	469	7	ABO67048	AbO67048 Klebsiell
19	105.5	39.7	372	5	ABG80418	Abg80418 Haemophil
20	105.5	39.7	372	7	ABO23507	AbO23507 Haemophil
21	104.5	39.3	35	3	AAY82571	Aay82571 Copolymer
22	104	39.1	100	3	AAY98499	Aay98499 Peptide #
23	104	39.1	100	3	AAY59044	Aay59044 Amino aci
24	104	39.1	100	4	AAB45852	Aab45852 Nucleic a
25	104	39.1	100	4	AAU04289	Aau04289 Poly-Lys-

ALIGNMENTS

RESULT 1

AAY82573

ID AAY82573 standard; peptide; 56 AA.

XX AC AAY82573;

DT 28-JUL-2000 (first entry)

XX

DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:3.

XX

Copolymer; molecular weight marker; TV-marker; immune disease; glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective; osteopathic; immunosuppressive; antithyroid; antiinflammatory; antidiabetic; thyronimetic; haemostatic; antipsoriatic; dermatological; antianemic; immunosuppressive; demyelinating disease; osteoarthritis; inflammatory condition; multiple sclerosis; rheumatoid arthritis; Crohn's disease; chronic immune thrombocytopaenia purpura; colitis; diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis; Hashimoto's disease; idiopathic myxoedema; myasthenia gravis; pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

OS

XX WO200018794-A1.

XX

PD 06-APR-2000.

XX

PF 24-SEP-1999; 99WO-US022402.

XX

PR 25-SEP-1999; 98US-0101693P.

XX

PA (YEDA) YEDA RES & DEV CO LTD.

PA (TEVA-) TEVA PHARM USA INC.

XX

PI Gad A, Lis D;

XX

DR WPI; 2000-317499/27.

XX

PT Copolymer 1 related polypeptides used as molecular weight markers for glatiramer acetate and for treatment and prevention of immune diseases.

XX

PS Claim 10; Page 14; 72pp; English.

XX

CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular weight TV-marker polypeptides from the present invention. The present invention describes polypeptides (I) for determining the molecular weight of a copolymer (CP), which has an identified molecular weight and an amino acid composition corresponding to the copolymer. The polypeptides

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 72.962 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-4
Perfect score: 313
Sequence: 1 AKYAKKEKAYAKKAEAK.....EAKYKAEAKAAKAEAYEA 66

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	DB	ID	Description
1	313	100.0	66	3	AAy82574	Copolymer	AAy82574 Copolymer
2	282.5	90.3	77	3	AAy82575	Copolymer	AAy82575 Copolymer
3	268	85.6	86	3	AAy82576	Copolymer	AAy82576 Copolymer
4	228.5	73.0	109	3	AAy82577	Copolymer	AAy82577 Copolymer
5	198	63.3	56	3	AAy82573	Copolymer	AAy82573 Copolymer
6	131	41.9	421	6	ABU28559	Protein e	ABU28559 Protein e
7	130	41.5	323	6	ABU31397	Protein e	ABU31397 Protein e
8	130	41.5	428	6	ABU27824	Protein e	ABU27824 Protein e
9	130	41.5	469	7	ABO67048	Klebsiell	ABO67048 Klebsiell
10	129.5	41.4	154	2	AAK06445	Recombina	AAK06445 Recombina
11	128	40.9	100	3	AAy98499	Peptide #	AAy98499 Peptide #
12	128	40.9	100	3	AAy59044	Amino aci	AAy59044 Amino aci
13	128	40.9	100	4	AAK45852	Nucleic a	AAK45852 Nucleic a
14	128	40.9	100	4	AAU04289	Poly-Lys-	AAU04289 Poly-Lys-
15	126.5	40.4	389	6	ABU39221	Protein e	ABU39221 Protein e
16	124.5	39.8	407	6	ABU47123	Protein e	ABU47123 Protein e
17	122.5	39.1	372	6	ABU40185	Protein e	ABU40185 Protein e
18	122.5	39.1	376	6	ABU47848	Protein e	ABU47848 Protein e
19	122	39.0	106	2	AAK06446	Recombina	AAK06446 Recombina
20	121	38.7	387	6	ABU44721	Protein e	ABU44721 Protein e
21	120.5	38.5	45	3	AAy82572	Copolymer	AAy82572 Copolymer
22	119.5	38.2	347	6	ABU38313	Protein e	ABU38313 Protein e
23	119.5	38.2	347	6	ABJ18771	Pseudomon	ABJ18771 Pseudomon
24	119.5	38.2	407	7	ABO80835	Pseudomon	ABO80835 Pseudomon
25	117.5	37.5	372	5	ABG80418	Haemophil	ABG80418 Haemophil

26	117.5	37.5	372	7	ABO23507	Haemophil	ABO23507 Haemophil
27	117	37.4	357	6	ABM67869	Photorhab	ABM67869 Photorhab
28	117	37.4	388	6	ABU50266	Protein e	ABU50266 Protein e
29	115	36.7	336	6	ABU42038	Protein e	ABU42038 Protein e
30	114.5	36.6	361	7	ADP05105	Bacterial	ADP05105 Bacterial
31	114	36.4	452	6	ABP57088	Mouse MKP	ABP57088 Mouse MKP
32	113.5	36.3	214	6	ADA33882	Acinetoba	ADA33882 Acinetoba
33	111.5	35.6	239	5	ABB49123	Listeria	ABB49123 Listeria
34	111.5	35.6	239	6	ABU32619	Protein e	ABU32619 Protein e
35	110.5	35.3	205	3	AAAB20575	Mycobacte	AAAB20575 Mycobacte
36	109.5	35.0	80	5	ABG71044	Tumour ne	ABG71044 Tumour ne
37	108.5	34.7	198	4	ABE95499	Human pro	ABE95499 Human pro
38	108.5	34.7	279	4	AAU03592	Human DNA	AAU03592 Human DNA
39	108.5	34.7	467	4	AAAB94309	Human pol	AAAB94309 Human pol
40	108.5	34.7	467	5	ABP69558	Human pol	ABP69558 Human pol
41	106	33.9	214	2	AAy34055	M. tuberc	AAy34055 M. tuberc
42	106	33.9	214	2	AAy57353	Protein e	AAy57353 Protein e
43	106	33.9	214	6	ABU34623	Protein e	ABU34623 Protein e
44	106	33.9	214	6	ABU36893	Protein e	ABU36893 Protein e
45	105.5	33.7	35	3	AAy82571	Copolymer	AAy82571 Copolymer

ALIGNMENTS

RESULT 1
AAy82574
ID AAy82574 standard; peptide; 66 AA.

XX AAy82574;

XX 28-JUL-2000 (first entry)

XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:4.

Copolymer; molecular weight marker; TV-marker; immune disease;
glutarimer acetate; autoimmune disease; antiarthritic; neuroprotective;
osteopathic; immunosuppressive; antithyroid; antiinflammatory;
antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
inflammatory condition; multiple sclerosis; rheumatoid arthritis;
Crohn's disease; chronic immune thrombocytopenia purpura; colitis;
diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.

XX WO200018794-A1.

XX 06-APR-2000.

XX 24-SEP-1999; 99WO-US022402.

XX 25-SEP-1998; 98US-0101693P.

XX (YEDA) YEDA RES & DEV CO LTD.
(TEVA-) TEVA PHARM USA INC.

XX Gad A, Lis D;

XX WPI; 2000-317499/27.

XX Copolymer 1 related polypeptides used as molecular weight markers for
glutarimer acetate and for treatment and prevention of immune diseases.

XX Claim 10; Page 14; 72pp; English.

XX AAy82571 to AAy82577 represent specifically claimed copolymer molecular
weight TV-marker polypeptides from the present invention. The present
invention describes polypeptides (I) for determining the molecular weight
of a copolymer (CP), which has an identified molecular weight and an
amino acid composition corresponding to the copolymer. The polypeptides

[illegible]

QY 49 KYKAEAAKAAKEAAEYA 66
| | | | | : | | | | : | |
Db 194 KKAEAAEAEEAAAEARKKA 211

RESULT 7
ABU31397
ID ABU31397 standard; protein; 323 AA.
AC ABU31397;
XX
DT 19-JUN-2003 (first entry)
DE Protein encoded by Prokaryotic essential gene #16924.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Klebsiella pneumoniae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA35267.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 59321; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 323 AA;

Query Match 41.5%; Score 130; DB 6; Length 323;
Best Local Similarity 57.7%; Pred. No. 4.6e-05;
Matches 41; Conservative 6; Mismatches 18; Indels 6; Gaps 3;

QY 1 AKKYAKKEKAYAKA---KKAFAKAAKAKAAEAKKYA--KAAKAEKKEVAAA--AKYKAA 54
| | | | | : | | | | | | | | | | : | | | | | | | | | |
Db 80 AKQAEAAEAAKAAAEAKQKAEAAAKAAQQAQAEKKAQQAQAAEAAEAAEAKAA 139
| | | | | : | | | | | | | | | | : | | | | | | | | | |
QY 55 AKAAAEKAAEYE 65
| | | | | : | | | | | | | | | | : | | | | | | | | | |
Db 140 QKAAAEKAAAE 150

RESULT 8
ABU27824
ID ABU27824 standard; protein; 428 AA.
XX
AC ABU27824;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #13351.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Enterobacter cloacae.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA31694.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 55748; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 428 AA;

Query Match 41.5%; Score 130; DB 6; Length 428;
 Best Local Similarity 56.1%; Pred. No. 6.2e-05;
 Matches 37; Conservative 11; Mismatches 16; Indels 2; Gaps 2;
 1 AKKYAKKEKAYAKAKAEAKAAKAAKAEAKKAYAKAEAKKAYAKAEAKKAAKAAK 60
 163 AAEAKKAAADAQ-KKAEAEAKKAAADAQKAE-AEAKKAAADAQKAEAEAKKAAQ 220
 61 EAAYEA 66
 221 EAEKKA 226

RESULT 9
 ABO67048
 ID ABO67048 standard; protein; 469 AA.
 XX
 AC ABO67048;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 13565.
 XX
 DE Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 US6610836-B1.
 PN
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA Breton GL, Osborne M;
 XX WPI; 2003-895346/82.
 DR N-PSDB; ABD00619.
 XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 XX
 PS Disclosure; SEQ ID NO 13565; 932pp; English.
 XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression

CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 469 AA;

Query Match 41.5%; Score 130; DB 7; Length 469;
 Best Local Similarity 57.7%; Pred. No. 6.9e-05;
 Matches 41; Conservative 6; Mismatches 18; Indels 6; Gaps 3;
 1 AKKYAKKEKAYAKA---KKAERAKAKAKAEAKKVA--KAAKAEKKEYAAAE-AKYKAA 54
 230 AKQAEAEAKAAAEAKQKAEAAAKAAQQAQAEKKAQQAQAEAKAAAEKAAAEKAA 289
 55 AKAAAEKAAVE 65
 290 QKAAAEKAAAE 300

RESULT 10
 AAR06445
 ID AAR06445 standard; protein; 154 AA.
 XX
 AC AAR06445;
 XX
 DT 25-MAR-2003 (revised)
 DT 03-JAN-1991 (first entry)
 XX
 DE Recombinant copolymer 1-77, myelin basic protein analogue.
 XX
 KW Recombinant copolymer 1; COP-1-77; myelin basic protein; MBP;
 KW immunological activity; autoimmune encephalomyelitis; multiple sclerosis.
 XX
 OS Synthetic.
 XX
 PN EP383620-A.
 XX
 PD 22-AUG-1990.
 XX
 PF 16-FEB-1990; 90EP-00301700.
 XX
 PR 17-FEB-1989; 89US-00312541.
 PR 07-FEB-1990; 90US-00473845.
 XX
 PA (REPK) REPLICEN CORP.
 XX
 PI Cook KS;
 XX
 DR WPI; 1990-255848/34.
 DR N-PSDB; AAQ05664.
 XX
 PT Producing genes encoding random polymers of aminoacid(s) - for producing
 PT recombinant polypeptide(s) with biological and/or immunological activity.
 XX
 PS Disclosure; Fig 11; 25pp; English.

CC To improve the expression of rCOP-1 polypeptides in *E. coli*, genes coding
 CC for rCOP-1-77 were subcloned from pREV 2.1 to pBG3-2deltaN (deposit: 20-
 CC NOV-1984 US4691009, NRRL B-15910), a plasmid used to express Protein A.
 CC The resulting plasmids encode fusion proteins consisting of beta-
 CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue
 CC occurs between the Protein A and rCOP-1 sequences, originating from the
 CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved
 CC from the fusion protein. rCOP-1-77 contains oligonucleotide duplexes
 CC incooding the following segments: YKK, EAE, KAK, AAK, and AAA. The N-
 CC terminal alanine residue is left behind following CNBr cleavage of the
 CC fusion protein. The product prevents or arrests experimental autoimmune
 CC encephalomyelitis. They are used to prevent, arrest or control a
 CC demyelinating disorder, e.g. multiple sclerosis. They may also be used as
 CC additives to hair care products to confer beneficial effects on damaged
 CC hair or as supplements for diets deficient in certain amino acids. See
 CC also AAQ05665. (Updated on 25-MAR-2003 to correct PA field.)
 XX


```

Query Match      40.9%; Score 128; DB 3; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

OY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEYAAAEAKYKAEA-AK 56
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61

OY 57 AAKEAA 63
    |||||
DB 62 AKAKAKA 68

RESULT 13
AAB45852
ID AAB45852 standard; protein; 100 AA.
XX
AC AAB45852;
XX
DT 21-MAR-2001 (first entry)
XX
DE Nucleic acid transporter system peptide ligand SEQ ID NO 64.
XX
KW Nucleic acid delivery; nucleic acid transporter system; hormone; enzyme;
KW growth factor; clotting factor; apolipoprotein; receptor; drug; oncogene;
KW tumor antigen; tumor suppressor; viral antigen; parasitic antigen;
KW bacterial antigen.
XX
OS Unidentified.
XX
PN US6150168-A.
XX
PD 21-NOV-2000.
XX
PF 05-JUN-1995; 95US-00460971.
XX
PR 20-MAR-1992; 92US-00855389.
XX
PR 19-MAR-1993; 93MO-US002725.
XX
PR 14-DEC-1993; 93US-00167641.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Gottchalk S, Sparrow J, Cristiano RJ, Smith LC, Woo SLC;
XX WPI; 2001-049093/06.
XX
PT Nucleic acid transporter system for delivering nucleic acid into a cell,
PT useful for delivering proteins and polypeptides to cells, including
PT growth factors, enzymes, hormones, and tumor suppressors.
XX
PS Disclosure; Col 125-126; 105pp; English.
XX
CC This invention describes a novel system (I) for delivering a nucleic acid
CC to a cell, comprising a binding complex comprising a ligand binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC surface ligand, and a second binding complex comprising a second binding
CC molecule noncovalently bound to a nucleic acid and covalently linked to a
CC nuclear ligand. The complexes are simultaneously bound to the nucleic
CC acid. The nucleic acid transporter system can also be used in a method
CC for the in vivo targeting of the insertion of DNA into a cell. It can
CC also be used in processes for producing transformed cell lines. The
CC system can be used to deliver a variety of proteins and polypeptides,
CC such as hormones, growth factors, enzymes, clotting factors,
CC apolipoproteins, receptors, drugs, oncogenes, tumor antigens, tumor
CC suppressors, viral antigens, parasitic antigens, and bacterial antigens.
CC The transporter system uses lysis agents to overcome the problems of
CC endosomal/lysosomal degradation seen with prior art systems
XX
SQ Sequence 100 AA;

Query Match      40.9%; Score 128; DB 4; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

OY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEYAAAEAKYKAEA-AK 56
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61

OY 57 AAKEAA 63
    |||||
DB 62 AKAKAKA 68

RESULT 14
AAU04289
ID AAU04289 standard; peptide; 100 AA.
XX
AC AAU04289;
XX
DT 23-OCT-2001 (first entry)
XX
DE Poly-Lys-Ala used in nucleic acid transporter system.
XX
KW Nucleic acid transport; cytosin; ligand; lysis agent; spacer molecule;
KW gene therapy; hepatocyte; muscle; bone forming cell.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 3..100
FT /note= "Lys-Ala in positions 3-100 may be present or
FT absent"
XX
PN US6177554-B1.
XX
PD 23-JAN-2001.
XX
PF 05-JUN-1995; 95US-00462040.
XX
PR 20-MAR-1992; 92US-00855389.
XX
PR 19-MAR-1993; 93MO-US002725.
XX
PR 14-DEC-1993; 93US-00167641.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
PI Woo SLC, Smith LC, Cristiano RJ, Gottchalk S, Sparrow J;
XX WPI; 2001-365933/38.
XX
PT Nucleic acid transport system, useful for creating transgenic animals for
PT assessing human disease such as cancer in an animal model.
XX
PS Disclosure; Col 131; 111pp; English.
XX
CC The sequence represents poly-Lys-Ala, used to bind nucleic acid in a
CC nucleic acid transporter system. The nucleic acid transporter system uses
CC nucleic acid binding complexes containing surface ligands which are
CC capable of binding to a cell surface receptor and entering the cell
CC through cytosin. The compounds of the invention are either ligands,
CC binding molecules (surface ligands), lysis agents, spacer molecules or
CC their intermediates. The ligands, binding molecules, lysis agents and
CC spacer molecules are used in nucleic acid transporter systems to deliver
CC nucleic acid into specific cells e.g. in gene therapy to deliver nucleic
CC acid into hepatocytes, muscle cells or bone forming cells
XX
SQ Sequence 100 AA;

Query Match      40.9%; Score 128; DB 4; Length 100;
Best Local Similarity 62.7%; Pred. No. 2.1e-05;
Matches 42; Conservative 5; Mismatches 16; Indels 4; Gaps 4;

OY 1 AKKYAK-KEKAYAK-KAEAKAAKAKAEAKKYAKA-AKAEKKEYAAAEAKYKAEA-AK 56
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61

```


QY 57 AAKKAA 63
Db 62 AAKAKA 68

RESULT 15

ABU39221
ID ABU39221 standard; protein; 389 AA.

XX AC ABU39221;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #24748.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Pasteurella multocida.

XX PN WO20027183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.

XX DR N-PSDB; ACA43091.

XX PT New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX PS Claim 25; SEQ ID NO 67145; 1766pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 389 AA;

Query Match 40.4%; Score 126.5; DB 6; Length 389;
Best Local Similarity 52.0%; Pred. No. 0.00012;
Matches 39; Conservative 9; Mismatches 12; Indels 15; Gaps 3;
QY 1 AKKYAKKAKAVAKAK-----KAEAKA---AKKAKAEAKKYAKAKAEKKEKYAAAEAKYK 51
Db 189 AKRKAEEKAKAEAEAKAKVKEKAKAEAEAKVKEKAKAEAEAKAKAEAKAKAEKAK 51
QY 52 AFAAKAAAKEAAAYEA 66
Db 243 AEKAKADAEEAAQORKA 257

Search completed: December 14, 2004, 06:01:13
Job time : 73.962 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:52:22 ; Search time 163.747 Seconds
(without alignments)
143.965 Million cell updates/sec

Title: US-10-792-311-4
Perfect score: 313
Sequence: 1 AKKYAKKEKAYAKAKAEAK.....EAKYKAEAKAAKAEAYEA 66

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	313	100.0	66	9 US-09-816-989A-4	Sequence 4, Appli
2	282.5	90.3	77	9 US-09-816-989A-5	Sequence 5, Appli
3	268	85.6	86	9 US-09-816-989A-6	Sequence 6, Appli
4	228.5	73.0	109	9 US-09-816-989A-7	Sequence 7, Appli
5	198	63.3	56	9 US-09-816-989A-3	Sequence 3, Appli
6	131	41.9	421	15 US-10-282-122A-56483	Sequence 56483, A
7	130	41.5	323	15 US-10-282-122A-59321	Sequence 59321, A
8	130	41.5	428	15 US-10-282-122A-55748	Sequence 55748, A
9	126.5	40.4	389	15 US-10-282-122A-67145	Sequence 67145, A
10	124.5	39.8	407	15 US-10-282-122A-75047	Sequence 75047, A
11	122.5	39.1	372	15 US-10-282-122A-68109	Sequence 68109, A
12	122.5	39.1	376	15 US-10-282-122A-75772	Sequence 75772, A
13	121	38.7	367	15 US-10-282-122A-72645	Sequence 72645, A

14	120.5	38.5	45	9 US-09-816-989A-2	Sequence 2, Appli
15	119.5	38.2	347	14 US-10-127-032-120	Sequence 120, App
16	119.5	38.2	347	15 US-10-282-122A-66237	Sequence 66237, A
17	117.5	37.5	372	10 US-09-820-843A-8	Sequence 8, Appli
18	117.5	37.5	372	16 US-10-467-421-16	Sequence 16, Appli
19	117	37.4	388	15 US-10-282-122A-78190	Sequence 78190, A
20	115	36.7	336	15 US-10-282-122A-69862	Sequence 69862, A
21	114	36.4	452	14 US-10-184-832-5	Sequence 5, Appli
22	113	36.1	272	17 US-10-739-930-10710	Sequence 10710, A
23	111.5	35.6	239	15 US-10-282-122A-60543	Sequence 60543, A
24	109.5	35.0	63	16 US-10-667-004-20	Sequence 20, Appli
25	108.5	34.7	84	15 US-10-424-599-269191	Sequence 269191, A
26	108.5	34.7	279	14 US-10-181-071-7	Sequence 7, Appli
27	106	33.9	214	15 US-10-229-567-27	Sequence 27, Appli
28	106	33.9	214	15 US-10-282-122A-62547	Sequence 62547, A
29	106	33.9	214	15 US-10-282-122A-64817	Sequence 64817, A
30	105.5	33.7	35	9 US-09-816-989A-1	Sequence 1, Appli
31	105.5	33.7	212	15 US-10-282-122A-61735	Sequence 61735, A
32	105.5	33.7	885	17 US-10-425-115-198446	Sequence 198446, A
33	105	33.5	183	16 US-10-767-701-61221	Sequence 61221, A
34	105	33.5	448	15 US-10-282-122A-45264	Sequence 45264, A
35	104.5	33.4	223	13 US-10-051-643-201	Sequence 201, App
36	104.5	33.4	223	14 US-10-205-979-52	Sequence 52, Appli
37	104.5	33.4	240	17 US-10-739-930-9712	Sequence 9712, Ap
38	104.5	33.4	507	16 US-10-726-692-50	Sequence 50, Appli
39	104.5	33.4	717	16 US-10-726-692-36	Sequence 36, Appli
40	104	33.2	578	16 US-10-437-963-119818	Sequence 119818, A
41	103.5	33.1	827	16 US-10-437-963-152005	Sequence 152005, A
42	103	32.9	114	16 US-10-767-701-53572	Sequence 53572, A
43	103	32.9	236	17 US-10-425-115-188164	Sequence 188164, A
44	102.5	32.7	433	15 US-10-390-472-2	Sequence 2, Appli
45	102	32.6	139	15 US-10-282-122A-60257	Sequence 60257, A

ALIGNMENTS

RESULT 1

US-09-816-989A-4
; Sequence 4, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; PRIOR FILING DATE: 2001-03-23
; PRIOR FILING DATE: 1998-09-25
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-4

Query Match	100.0%	Score 313;	DB 9;	Length 66;
Best Local Similarity	100.0%	Pred. No. 7.6e-22;	Indels	0;
Matches	66;	Conservative	0;	Mismatches

Qy	1	AKKYAKKEKAYAKAKAEAKAAKAEAKYKAAKAEKKEKYAAAEAKYKAEAAKAAK 60
Db	1	AKKYAKKEKAYAKAKAEAKAAKAEAKYKAAKAEKKEKYAAAEAKYKAEAAKAAK 60
Qy	61	EAAVEA 66


```

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75047
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75047

Query Match          39.8%; Score 124.5; DB 15; Length 407;
Best Local Similarity 47.8%; Pred. No. 0.00097;
Matches 43; Conservative 6; Mismatches 10; Indels 31; Gaps 4;

QY   1 AKKYAKKEKAYAKA-----KAAEAKA-----KAAEAKKYAKAAEAKEY 43
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db   189 AKKAAEAAKAAAEAKKAEEAKAAEAKAAEAKKKADAAEAKAAEAKKADA 244
      ||| | | | | | | | | | | | | | | | | | | | | | | | | |

QY   44 AAEAKEYKAEAA-----KAAAKEAA 63
      ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db   245 AAAEAKKADAAAACAAADAKKKAAAEKAA 274
      ||| | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-10-282-122A-68109
; Sequence 68109, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75772
; LENGTH: 376
; TYPE: PRT

RESULT 12
US-10-282-122A-75772
; Sequence 75772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75772
; LENGTH: 376
; TYPE: PRT

```

```

; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75047
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75047

Query Match          39.8%; Score 124.5; DB 15; Length 407;
Best Local Similarity 47.8%; Pred. No. 0.00097;
Matches 43; Conservative 6; Mismatches 10; Indels 31; Gaps 4;

QY   1 AKKYAKKEXAYAKA-----KKAEEAKA-----KAKAEKKYAKAAEAKEY 43
      ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   189 AKKAAEAAKAAEAEEKKAAEAKEAAKAAEAKEAAEKADAAEAKEAAKADAAA 244
      ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY   44 AAAEAKYKAEAA-----KAAAEKAA 63
      ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db   245 AAEEAKKADAAAAAACAADAADAKKAAAEKAA 274
      ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 11
US-10-282-122A-68109
; Sequence 68109, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75772
; LENGTH: 376
; TYPE: PRT

RESULT 12
US-10-282-122A-75772
; Sequence 75772, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75772
; LENGTH: 376
; TYPE: PRT

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 18.519 Seconds
(without alignments)
236.351 Million cell updates/sec

Title: US-10-792-311-4

Perfect score: 313

Sequence: 1 AKKYAKKEKAYAKAKKEAK.....EAKYKAEAAKAAKEAAVEA 66

Scoring table: BLOSUM62

scoring cable: BROSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 478139 segs. 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0% Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

```
Database: loaded packages:
1: /cqn2 6/ptodata/1/iaa/5A COMB.per:*
```

2: /cqn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cqn2_6/ptodata/1/iaa/6A_COMB.per:*

4: /cgn2 6/ptodata/1/iaa/6B_COMB.pep:*

5: /cgn2-6/ptodata/1/iaa/PCTUS COMB.pcp:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	313	100.0	66	4	US-09-405-743A-4		Sequence 4, Appli
2	282.5	90.3	77	4	US-09-405-743A-5		Sequence 5, Appli
3	268	85.6	86	4	US-09-405-743A-6		Sequence 6, Appli
4	228.5	73.0	109	4	US-09-405-743A-7		Sequence 7, Appli
5	198	63.3	56	4	US-09-405-743A-3		Sequence 3, Appli
6	130	41.5	469	4	US-09-489-039A-13565		Sequence 13565, A
7	128	40.9	100	2	US-08-460-890A-64		Sequence 64, Appl
8	128	40.9	100	3	US-08-167-641C-64		Sequence 64, Appl
9	128	40.9	100	3	US-08-460-971A-64		Sequence 64, Appl
10	128	40.9	100	3	US-08-462-040-64		Sequence 64, Appl
11	120.5	38.5	45	4	US-09-405-743A-2		Sequence 2, Appli
12	119.5	38.2	407	4	US-09-252-991A-29581		Sequence 29581, A
13	117.5	37.5	56	3	US-08-993-008A-6		Sequence 6, Appli
14	114.5	36.6	361	4	US-09-343-681A-5390		Sequence 5390, Ap
15	113.5	36.3	214	4	US-09-328-352-5169		Sequence 5169, Ap
16	107	34.2	48	3	US-08-993-008A-5		Sequence 5, Appli
17	106	33.9	214	3	US-09-041-889-27		Sequence 27, Appl
18	106	33.9	214	4	US-09-417-264-27		Sequence 27, Appl
19	105.5	33.7	35	4	US-09-405-743A-1		Sequence 1, Appli
20	105	33.5	1507	3	US-08-929-329-5		Sequence 5, Appli
21	104.5	33.4	223	3	US-09-095-855-201		Sequence 201, App
22	104.5	33.4	223	4	US-09-205-426-201		Sequence 201, App
23	102.5	32.7	433	1	US-08-346-849-2		Sequence 2, Appli
24	102.5	32.7	433	2	US-08-283-284A-2		Sequence 2, Appli
25	102.5	32.7	433	4	US-08-998-300-2		Sequence 2, Appli
26	102.5	32.7	468	4	US-09-328-352-6321		Sequence 6321, Ap
27	102	32.6	207	4	US-09-489-039A-13743		Sequence 13743, A


```

; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
;
US-09-405-743A-2

Query Match      38.5%; Score 120.5; DB 4; Length 45;
Best Local Similarity 71.1%; Pred. No. 2.3e-05;
Matches 33; Conservative 0; Mismatches 6; Indels 7; Gaps 2;

QY      29 AKKYAKAAKAE--KKEYAAAEAK-----YKAEAAKAAAKAEAAVEA 66
Db      1 AKKYAKKAAKAEKAKYKAAAEKAAKAAKYKAAAEKAAKAAAEAAVEA 45
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
US-09-252-991A-29581
; Sequence 29581, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-29581

Query Match      38.2%; Score 119.5; DB 4; Length 407;
Best Local Similarity 51.5%; Pred. No. 0.00027;
Matches 34; Conservative 7; Mismatches 22; Indels 3; Gaps 1;

QY      1 AKKYAK---KEKAYAKAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 57
Db      215 AKKRADEAKKAAEDAKKAAEDAKKAAEDAKKAAEDAKKAAEDAKKAAEDAKKAA 274
          ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
          ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY      58 AAKEAA 63
Db      275 AAKKAA 280
          ||::||
          ||::||

RESULT 13
US-08-993-008A-6
; Sequence 6, Application US/08993008A
; Patent No. 6153596
; GENERAL INFORMATION:
; APPLICANT: Liotta, Dennis C.
; APPLICANT: Petros, John A.
; APPLICANT: Wey, Shioh-Jyi
; APPLICANT: Karr, Joan F.
; APPLICANT: Pohl, Jan
; TITLE OF INVENTION: Polycationic Oligomers
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan
; STREET: 5370 Manhattan Circle, Suite 201

```

Db 233 KAAAOQ 238

RESULT 15

US-09-328-352-5169
; Sequence 5169, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5169
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5169

Query Match 36.3%; Score 113.5; DB 4; Length 214;
Best Local Similarity 51.4%; Pred. No. 0.00052;
Matches 37; Conservative 9; Mismatches 17; Indels 9; Gaps 4;
QY 1 AKKYAKKEKAYAKAKAEAKKAKAEAKKYAKAAK--AEKKEVA-----AAEAKYKAE 53
Db 55 AKRQAEADKKAAEAKR-QAEADKKA-AEAKRQAEADKKAAEAKRQAEADKKAAEAKRKA 112
QY 54 AAQAAAEKAAE 65
Db 113 AEKKAEEAKAE 124

Search completed: December 14, 2004, 05:50:17
Job time : 18.6618 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	150.5	41.1	421	2	JV0057	tola protein - Esc
2	148	40.4	394	2	F90725	membrane spanning
3	148	40.4	394	2	G85576	membrane spanning
4	145	39.9	376	2	AG0592	tola protein [impo
5	141.5	38.7	347	2	E83525	Tola protein PA097
6	141	38.5	372	2	G64064	outer membrane int
7	138	37.7	1701	2	T09127	probable erythrocy
8	136	37.2	210	2	A25550	histone H1 - sea u
9	134.5	36.7	219	2	E60110	repetitive protein
10	134.5	36.7	231	2	S59589	histone H1 - Chlam
11	134	36.6	1128	2	T30296	R27-2 protein - Tr
12	133	36.3	328	2	A43993	cytosolic repetiti
13	130.5	35.7	388	2	AC0138	Tola colicin impor
14	130	35.5	1390	2	S51364	sperm tail-specifi
15	129.5	35.4	214	2	G70673	probable hupB - My
16	128.5	35.1	211	2	A28100	histone H1-beta, e
17	127.5	34.8	344	2	S34153	mst101-1 protein -
18	127	34.7	217	2	A26721	histone H1-gamma,
19	123	33.6	248	1	HSURIP	histone H1, gonada
20	123	33.6	433	2	S25194	zuoitin - yeast (Sa
21	122.5	33.5	206	2	S09388	histone H1 - sea u
22	122.5	33.5	220	2	A28456	histone H1.10 - ch
23	122.5	33.5	311	2	T17698	hypothetical prote
24	122	33.3	243	2	AE1689	hypothetical prote
25	121.5	33.2	239	2	AE1317	hypothetical prote
26	121.5	33.2	384	2	B43592	outer membrane pro
27	120.5	32.9	284	2	T06241	histone H1 (clone
28	118.5	32.4	241	2	JN0748	histone H1-II - Vo
29	118	32.2	206	1	HST914	histone H1 - rainb

```

membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90725
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.-G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <HAY>
A;Cross-references: UNIPROT:Q8X965; GB:BA000007; PIDN:BAB34197.1; PID:g13360233; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: EC80774

Query Match 40.4%; Score 148; DB 2; Length 394;
Best Local Similarity 55.1%; Pred. No. 0.00033;
Matches 49; Conservative 7; Mismatches 21; Indels 12; Gaps 4;

QY 1 AKKYAKKEKAYA-----KKAEEAKKAEAKYKAAEAK-KKAKAEAKKYAKAA-----KA 49
Db 138 AAKAAADAKAKAADDKAAEEAKKAAADAKKKAEAAKAAAEAKKKAEAAALKKKA 197

QY 50 EKKEYAAAEAKYKAAKAAA-KEAAVEA 77
Db 198 EAAEAAAAAEAKKAAAEKAAADAKKAAAEKA 226

RESULT 3
G85576
membrane spanning protein Tola [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85576
R;Perna, L.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85576
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-394 <STO>
A;Cross-references: UNIPROT:Q8X965; GB:AE005174; NID:g12513672; PIDN:AAG55075.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: tola

Query Match 40.4%; Score 148; DB 2; Length 394;
Best Local Similarity 55.1%; Pred. No. 0.00033;
Matches 49; Conservative 7; Mismatches 21; Indels 12; Gaps 4;

QY 1 AKKYAKKEKAYA-----KKAEEAKKAEAKYKAAEAK-KKAKAEAKKYAKAA-----KA 49
Db 138 AAKAAADAKAKAADDKAAEEAKKAAADAKKKAEAAKAAAEAKKKAEAAALKKKA 197

QY 50 EKKEYAAAEAKYKAAKAAA-KEAAVEA 77
Db 198 EAAEAAAAAEARKKAAAEKAAADAKKAAEKA 226

RESULT 4
AG0592
tola protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain Ctr18
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0592
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

```

```

, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-376 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05209.1; PID:g16501979; GSPDB:GN00176
C;Genetics:
A;Gene: STY0793

Query Match 39.9%; Score 146; DB 2; Length 376;
Best Local Similarity 54.4%; Pred. No. 0.00044;
Matches 49; Conservative 7; Mismatches 14; Indels 20; Gaps 5;

QY 1 AKKYAKKEKAYAKAAKAAKAAKAYK-AAEAKKAKAAEAKKYAKAAKAAEKK-----EY 54
Db 158 AKKKAEEA--AKAAADAKKAAEAAKAAAEAKKAAEAAK--AAADAKKKADAEEAAK 213

QY 55 AAEAKYKAEAA-----KAAAKEAA 74
Db 214 AAEEAKKKADAAAKAAADAKKKAAAEKAA 243

RESULT 5
E83525
TolA protein PA0971 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83525
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Brj
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83525
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: UNIPROT:P50600; GB:AE004530; GB:AE004091; NID:g9946865; PIDN:AAG0436
A;Experimental source: strain PA01
C;Genetics:
A;Gene: tola; PA0971

Query Match 38.7%; Score 141.5; DB 2; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00089;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYAKAAKAAKAAKAEKA-----YKAAE-A 32
Db 120 ARKAAEQAAEAARAKKAAEAKKAAAEAKKAAEQKQADIAKKRAEDPAKKAAEDAKKKAEDA 179

QY 33 KKKAAEAKKYAKAAKAEKKEVAAAEAKYKAAAEKAAAEAA 74
Db 180 KKKAAEAEAKKA-AAEAKKKAAVEAKKKAASAAAAARAA 220

RESULT 6
G64064
outer membrane integrity protein tola - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C;Accession: G64064; JC5212
R;Gelschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630; PMID:7542800

```

A:Accession: G64064
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <TIG>
A:Cross-references: UNIPROT:P44678; GB:U32722; GB:L42023; NID:g1573348; PIDN:AAC22041.1;
R:Sen, K.; Sikkema, D.J.; Murphy, T.F.
Gene 178, 75-81, 1996
A:Title: Isolation and characterization of the Haemophilus influenzae tolQ, tolR, tolA
A:Reference number: JCS212; MUID:97080550; PMID:8921895
A:Accession: JCS212
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 'V', 2-47, 'A', 49-141, 'R', 143-164, 'P', 166-189, 'R', 191-202, 'A', 204-226, 'A', 228-
A:Cross-references: GB:U32470
A:Experimental source: strain 1479
A>Note: the authors translated the codon CGT for residue 190 as Ala
C:Genetics:
A:Gene: tolA
A:Start codon: GTG

Query Match 38.5%; Score 141; DB 2; Length 372;
Best Local Similarity 54.7%; Pred. No. 0.001;
Matches 47; Conservative 7; Mismatches 20; Indels 12; Gaps 4;

QY 1 AKYA-----KKEKAYAKAEAAK-----AEAKYKAEAKKAKAEAKKYAKAKAEK 51
DB 158 AKRLAAAKQAEFEAKAKAEIAAQKQAEAKAKLEAEAKAKAEAAK--AK-AEAEA 214
QY 52 KEYAAAEAKYKAEAKAKAAAEAAEAE 77
DB 215 KAKAAAEAKAKADAEAKAEAKRKA 240

RESULT 7
T09127
probable erythrocyte-binding protein MAEBL - Plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09127
R:Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A:Title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A:Reference number: Z16577; MUID:98115903; PMID:9448314
A:Accession: T09127
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1701 <KAP>
A:Cross-references: UNIPROT:O61164; EMBL:AF031886; NID:g2947227; PID:g2947228
A:Experimental source: subspecies yoelii; strain YM
C:Genetics:
A:Gene: maeb1
A:Introns: 62/1; 1648/1; 1674/2; 1697/1
C:Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 37.7%; Score 138; DB 2; Length 1701;
Best Local Similarity 54.4%; Pred. No. 0.0053;
Matches 43; Conservative 8; Mismatches 16; Indels 12; Gaps 4;

QY 5 AKKAYAKAEAKKAEAKKAEAKKAKA-----EAKYAKAA-KAEKKEA--- 55
DB 1224 AKKAEAAKAEERKKAEAKA-AKKALERKKSEAKKALERKKAEAKKAEKKAEAA 1282

QY 56 --AAEAKYKAEAAKAAKE 72
DB 1283 KKAEEKKKAEAAKAAEE 1301

RESULT 8
A25550
histone H1 - sea urchin (Lytechinus pictus)
C:Species: Lytechinus pictus (painted urchin)
C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C:Accession: A25550

R:Knowles, J.A.; Childs, G.J.
Nucleic Acids Res. 14, 8121-8133, 1986
A:Title: Comparison of the late H1 histone genes of the sea urchins Lytechinus pictus and
A:Reference number: A25550; MUID:87040778; PMID:3022245
A:Accession: A25550
A:Molecule type: DNA
A:Residues: 1-210 <KNO>
A:Cross-references: UNIPROT:P06144; GB:X04488; NID:g9616; PIDN:CAA28177.1; PID:g9617
C:Superfamily: histone H1
C:Keywords: chromosomal protein; DNA binding; nucleosome; nucleus

Query Match 37.2%; Score 136; DB 2; Length 210;
Best Local Similarity 53.3%; Pred. No. 0.0015;
Matches 40; Conservative 7; Mismatches 26; Indels 2; Gaps 2;

QY 2 KKYAKKERAYAKAEAKAEAKYKAAEAK-KKAEAKKYA-KAAKAEKKEAEEA 59
DB 96 KTEAQKARAAAKKLAAKKKEQEKKAATKARKEKLAAKAAKAAKVKKPAKAKK 155
QY 60 KYAEAAKAAAEAA 74
DB 156 PAKKAARKPAAKAA 170

RESULT 9
E60110
repetitive protein antigen 27 - Trypanosoma cruzi (fragments)
C:Species: Trypanosoma cruzi
C:Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 09-Jul-2004
C:Accession: E60110
R:Hoft, D.F.; Kim, K.S.; Otsu, K.; Moser, D.R.; Yost, W.J.; Blumin, J.H.; Donelson, J.E.;
Infect. Immun. 57, 1959-1967, 1989
A:Title: Trypanosoma cruzi expresses diverse repetitive protein antigens.
A:Reference number: A60110; MUID:89277508; PMID:2659529
A:Accession: E60110
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-219 <HOF>
A:Cross-references: UNIPROT:Q26947
C:Comment: This protein contains a series of tandem repeats, each fourteen residues in le
C:Superfamily: neurofilament triplet H protein
C:Keywords: tandem repeat

Query Match 36.7%; Score 134.5; DB 2; Length 219;
Best Local Similarity 50.7%; Pred. No. 0.002;
Matches 38; Conservative 8; Mismatches 28; Indels 1; Gaps 1;

QY 1 AKKYAKKERAYAKAEAKAEAKYKAAEAKKAEAKKAEAKKYAKAK-AEKKEAEEA 59
DB 70 ATKVAGDEKQAAEATKVAAEAKQKAEATKVAAEAKQKAEATKVAAEAKQKAEAA 129
QY 60 KYAEAAKAAAEAA 74
DB 130 TKVAEAGKQKAEAA 144

RESULT 10
S59589
histone H1 - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59589; S62122
R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Corneliuss, T.; Schmitt, R.
Curr. Genet. 28, 333-345, 1995
A:Title: The organization structure and regulatory elements of Chlamydomonas histone gene
A:Reference number: S59581; MUID:96120862; PMID:8590479
A:Accession: S59589
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-231 <FAB>
A:Cross-references: UNIPROT:Q39576; EMBL:U16726
A>Note: the authors did not translate the codon for residue 1
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Corneliuss, T.; Schmitt, R.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 86.4219 Seconds
(without alignments)
512.646 Million cell updates/sec

Title: US-10-792-311-5
 Perfect score: 366
 Sequence: 1 AKTYAKKEKAYAKAEKAAK.....EAKYKAEAAKAAKEAAVEA 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : UniProt_02:*
1: uniprot_sprot
2: uniprot_tr
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	152.5	41.7	41.3	2	Q7C2Q4		Q7C2q4 shigella fl
2	152.5	41.7	41.3	2	Q83SA1		Q83sa1 shigella fl
3	151	41.3	37.9	2	Q7WNF5		Q7wcn5 bordetella
4	150.5	41.1	42.1	1	TOLA_ECOLI		P19934 escherichia
5	150	41.0	37.3	2	Q7W477		Q7w477 bordetella
6	148	40.4	39.4	2	Q7AGI8		Q7ag18 escherichia
7	148	40.4	39.4	2	Q8X965		Q8x965 escherichia
8	148	40.4	40.7	2	Q8ZOT6		Q8zot6 salmonella
9	147.5	40.3	42.1	2	Q8FUT1		Q8fjt1 escherichia
10	146	39.9	37.6	2	Q8Z8C1		Q8z8c1 salmonella
11	145.5	39.8	39.5	2	Q937K4		Q937k4 erwinia chr
12	143	39.1	17.01	2	Q7RC08		Q7rc08 plasmodium
13	142.5	38.9	37.2	2	Q9WXX1		Q9wwx1 pseudomonas
14	142.5	38.9	37.2	2	Q8NNI6		Q8nni6 pseudomonas
15	142	38.8	39.5	2	Q6D7F3		Q6d7f3 erwinia car
16	141.5	38.7	34.7	1	TOLA_PSEAE		P50600 pseudomonas
17	141	38.5	37.2	1	TOLA_HAEIN		P44678 haemophilus
18	140	38.3	38.9	2	Q9CM70		Q9cm70 pasteurella
19	139.5	38.1	19.8	2	Q6MIU4		Q6miu4 bdellovibri
20	139.5	38.1	19.8	2	CAB80819		Caes0819 bdellovib
21	138.5	37.8	35.6	2	Q7N5T7		Q7n5t7 photorhabdu
22	138	37.7	23.2	2	Q39576		Q39576 chlamydomon
23	138	37.7	165.2	2	Q7KPE21		Q7kpe21 plasmodium
24	138	37.7	165.2	2	AAQ73456		AAq73456 plasmodiu
25	138	37.7	1680	2	Q7KPY9		Q7kpy9 plasmodium
26	138	37.7	1680	2	AAQ73455		AAq73455 plasmodiu
27	138	37.7	1680	2	AAQ73457		AAq73457 plasmodiu
28	138	37.7	1701	2	O61164		O61164 plasmodium
29	136	37.2	210	1	HI_LYTPI		P06144 lytechinus
30	136	37.2	71.3	2	O6CDX0		O6cdx0 yarrowia li
31	135.5	37.0	134.1	2	Q98KG7		Q98kg7 rhizobium l

RESULT 1

Q7C2Q4	PRELIMINARY;	PRT	413 AA.
Q7C2Q4;			
AC	Q7C2Q4;		
DT	05-JUL-2004	(TrEMBLrel. 27, Created)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
DE	Membrane spanning protein.		
GN	Name=tolA; OrderedLocusNames=S0571;		
OS	Shigella flexneri.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Shigella.		
OX	NCBI_TaxID=623;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=2457T;		
RX	MEDLINE=22590274; PubMed=12704152;		
RA	Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,		
RA	Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,		
RA	Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,		
RA	Schwartz D.C., Blattner F.R.;		
RT	"Complete genome sequence and comparative genomics of Shigella		
RT	flexneri serotype 2a strain 2457T.";		
RL	Infect. Immun. 71:2775-2786(2003).		
DR	EMBL; AE016979; AAP16075.1; -.		
DR	InterPro; IPR010528; TolA.		
DR	Pfam; PF06519; TolA; 1.		
SO	SEQUENCE 413 AA; 42355 MW; 93E10F2C5DE60D8 CRC64;		

Query Match 41.7%; Score 152.5; DB 2; Length 413;
Best Local Similarity 52.4%; Pred. No. 0.00077;

QY 1 AKTYAKKEKAYAKKA EKA KA AKA EA KAY KA AE AKA KA EAKYAK-----AAAEKK-E 53
:: :: : :: :: :: | | | | | | | | | | | | : :: :: :: |
pb 120 AFEAAKOA RI.KOKOAEVA AKA KA AD DA KA AFEAAKTAADAKKAEAEAAKAAEAEOAKAE 179

54 YAAAEAKYKAEAAKAAAKEAAVEA 77

Db 180 VAAALKKKAEAAEAAAAEARKK 203

RESULT. T 2

Q83SA1	PRELIMINARY; PRT; 413 AA.
ID	Q83SA1
AC	Q83SA1;
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Membrane spanning protein, required for outer membrane integrity.
GN	Name=tOra; OrderedLocusNames=Sf0558;
OS	Shigella flexneri.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Shigella.


```

RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffly binding.
DR Pfam; PF05424; Duffly binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 39.1%; Score 143; DB 2; Length 1701;
Best Local Similarity 55.3%; Pred. No. 0.012;
Matches 47; Conservative 7; Mismatches 17; Indels 14; Gaps 5;

QY 1 AKKYAKKEKA-YAKKAEKAAKAEAKYAKAEAKKAKA-----EAKYAKAAK---AEK 51
Db 1270 AKKAEKKKAAKAAKAEKKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEK 1328
QY 52 KEVAAA-----EAKYAKAEKAAKAAKE 72
Db 1329 KKAEEAKKAEKKKAEKAAKAAKE 1353

RESULT 13
Q9WXX1 PRELIMINARY; PRT; 372 AA.
AC Q9WXX1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE Biopolymer transport protein TolA.
DE TolA protein.
GN Name:tolA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=961981174; PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope.";
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida.";
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; ILR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.

```

```

DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TolA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 38.9%; Score 142.5; DB 2; Length 372;
Best Local Similarity 46.9%; Pred. No. 0.0039;
Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY 1 AKKYAKKEKA-----YAKKAEKAAKKA---EAKYKAAEAKKAKAKAKYA--- 44
Db 148 AKKAEKQADTAKKAEDEAKKAEKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAED 207
QY 45 ---KAAKAEKYYAAAEAKYK---AEAKAAKAAEAYEA 77
Db 208 AKKAAAEADAKKAAEAKKAAEAKKAAADAKKKAQEAARKA 245

RESULT 14
Q88NI6 PRELIMINARY; PRT; 372 AA.
AC Q88NI6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biopolymer transport protein TolA.
GN Name:tolA; OrderedLocustNames=PP1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohseil J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AS016778; AAN66845.1; -.
DR HSSP; P50600; ILR0.
DR TIGR; PP1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TolA.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 38.9%; Score 142.5; DB 2; Length 372;
Best Local Similarity 46.9%; Pred. No. 0.0039;

```

Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;
QY 1 AKKYAKKEKA-----YAKKAEKAAKKA---EAKYKAAEAKKKAKAEAKKYA--- 44
DB 148 AKKAAEKQADIAKKAEDEAKKAAEEAKKAAAEADAKKAAEEAKKKAED 207
QY 45 ---KAAKAEKYYAAAEAKYK--AEAKKAAAEAYEA 77
DB 208 AKKGAADAEAKKAAEAEAKKAAADAKKKAQEAARKA 245

RESULT 15

Q6D7F3

ID Q6D7F3 PRELIMINARY; PRT; 395 AA.

AC Q6D7F3;

DT 01-OCT-2004 (TReMBLrel. 28, Created)

DT 01-OCT-2004 (TReMBLrel. 28, Last sequence update)

DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)

DE Tola protein.

GN Name=tola; ORFNames=ECAL372;

OS Erwinia carotovora subsp. atroseptica SCRI1043.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OX NCBI_TaxID=218491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SCRI1043;

RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,

RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,

RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,

RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,

RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,

RA Salmond G.P.C., Birch P.R.J., Barrell B.G., Parkhill J., Toth I.K.;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX950851; CAG74282.1; -

SQ SEQUENCE 395 AA; 41698 MW; 2DB1DAA64E984199 CRC64;

Query Match

Best Local Similarity 38.8%; Score 142; DB 2; Length 395;

Matches 46; Conservative 51.1%; Pred.No.0.0044;

Matches 46; Conservative 9; Mismatches 17; Indels 18; Gaps 4;

QY 5 AKKEKAYAKKAEKAAKAEKAYK----AAEAKKKAKAEAKKYAKAKA-----EKKE 53

DB 134 AQQAKEQQKQAEAAAKAKAEAEFQQTAAADARKKAEDAEAKKQAVAAAAAKKQAEKE 193

QY 54 YAAAEAKY-----KAEAAKAAKAEAYEA 77

DB 194 KAAEAAKQKAAETAKAEAAKAA--EAKKEA 222

Search completed: December 14, 2004, 06:10:18

Job time : 86.4219 secs

This Page Blank (uspto)

CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 77 AA;

Query Match 100.0%; Score 366; DB 3; Length 77;
 Best Local Similarity 100.0%; Pred. No. 7.3e-27;
 Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKAYAKAEAKKAEKAYAAAEAK 60
 Db 1 AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKAYAKAEAKKAEKAYAAAEAK 60

QY 61 YKAEAAKAAKAEAAEYEA 77

Db 61 YKAEAAKAAKAEAAEYEA 77

RESULT 2
 AAY82576
 ID AAY82576 standard; peptide; 86 AA.

XX AAY82576;
 XX
 AC
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.
 XX
 KW Copolymer; molecular weight marker; TV-marker; immune disease;
 KW Glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.
 OS
 XX WO200018794-A1.
 PN
 XX
 XX
 PD 06-APR-2000.
 PF
 XX 24-SEP-1999; 99WO-US022402.
 XX
 XX 25-SEP-1998; 98US-0101693P.
 PR
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA (TEVA-) TEVA PHARM USA INC.
 XX
 XX Gad A, Lis D;
 PI
 XX WPI; 2000-317499/27.
 DR
 XX
 XX Copolymer 1 related polypeptides used as molecular weight markers for
 PT Glatiramer acetate and for treatment and prevention of immune diseases.

XX Claim 10; Page 14; 72pp; English.
 PS
 XX
 CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular
 CC weight TV-marker polypeptides from the present invention. The present
 CC invention describes polypeptides (I) for determining the molecular weight
 CC of a copolymer (CP), which has an identified molecular weight and an
 CC amino acid composition corresponding to the copolymer. The polypeptides
 CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopaenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 86 AA;

Query Match 96.0%; Score 351.5; DB 3; Length 86;
 Best Local Similarity 89.5%; Pred. No. 1.8e-25;
 Matches 77; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKAYAKAEAKKAEKAYAAAEAK 60
 Db 1 AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKAYAKAEAKKAEKAYAAAEAK 60

QY 61 -----YKAEAAKAAKAEAAEYEA 77

Db 61 YKAEAAKAYKAEAKAAKAEAAEYEA 86

RESULT 3
 AAY82577
 ID AAY82577 standard; peptide; 109 AA.

XX AAY82577;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:7.
 XX
 KW Copolymer; molecular weight marker; TV-marker; immune disease;
 KW Glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.

XX Unidentified.
 OS
 XX WO200018794-A1.
 PN
 XX
 XX 06-APR-2000.
 PD
 XX 24-SEP-1999; 99WO-US022402.
 PF
 XX 25-SEP-1998; 98US-0101693P.
 PR
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA

CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 387 AA;

Query Match 40.0%; Score 146.5; DB 6; Length 387;
 Best Local Similarity 58.8%; Pred. No. 9.2e-06;
 Matches 51; Conservative 7; Mismatches 16; Indels 13; Gaps 6;
 QY 1 AKKYAKKEKAYAKAEAKKAEAKAYK-AAEAKKKAKAEAKKYA----KAAKAEKKEYA 55
 Db 141 AKKKAPEA--AKAADAKKKAEAEVKAADAKKKAEAEAAKAAADAKKKAEAEAAK-A 197
 QY 56 ABAKAYK--AEAKAA--KAAAYEA 77
 Db 198 ABAKKAEEAEAKAAADAKKADAEEA 224

RESULT 9
 ABU47848
 ID ABU47848 standard; protein; 376 AA.
 XX
 AC ABU47848;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #33375.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Salmomella typhi.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Hasebeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA51718.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 75772; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 376 AA;

Query Match 39.9%; Score 146; DB 6; Length 376;
 Best Local Similarity 54.4%; Pred. No. 9.9e-06;
 Matches 49; Conservative 7; Mismatches 14; Indels 20; Gaps 5;
 QY 1 AKKYAKKEKAYAKAEAKKAEAKAYK-AAEAKKKAEAEAKKYAKAEAKAEKK-----EY 54
 Db 158 AKKKAPEA--AKAADAKKKAEAEAAKAAAEAKKKAEAEAAK--AAADAKKKADAEEAAK 213
 QY 55 ABAEAKYKAEAA-----KAAAKEAA 74
 Db 214 ABAEAKKKADAAAKAAADAKKKAEAKAA 243

RESULT 10
 AAY98499
 ID AAY98499 standard; peptide; 100 AA.
 XX
 AC AAY98499;
 XX
 DT 31-JUL-2000 (first entry)
 XX
 DE Peptide #10 used in nucleic acid transporter system.
 XX
 KW Transporter system; nucleic acid delivery; gene therapy; cancer;
 KW carcinogenesis; cardiovascular disease; infection.
 XX
 OS Synthetic.
 XX
 PN US6033884-A.
 XX
 PD 07-MAR-2000.
 XX
 PF 14-DEC-1993; 93US-00167641.
 XX
 PR 20-MAR-1992; 92US-00855389.
 PR 19-MAR-1993; 93WO-US002725.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:52:22 ; Search time 191.038 Seconds
(without alignments)
143.965 Million cell updates/sec

Title: US-10-792-311-5
Perfect score: 366
Sequence: 1 AKKYAKKEKAYAKAEKAAK.....EAKYKAEAKAAKAEAYEA 77

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues
Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match	Length DB ID	Description
1	366	100.0	77 9	US-09-816-989A-5 Sequence 5, Appli
2	351.5	96.0	86 9	US-09-816-989A-6 Sequence 6, Appli
3	289	79.0	109 9	US-09-816-989A-7 Sequence 7, Appli
4	282.5	77.2	66 9	US-09-816-989A-4 Sequence 4, Appli
5	245.5	67.1	56 9	US-09-816-989A-3 Sequence 3, Appli
6	150.5	41.1	421 15	US-10-282-122A-56483 Sequence 56483, A
7	148	40.0	407 15	US-10-282-122A-75047 Sequence 75047, A
8	146.5	40.0	387 15	US-10-282-122A-72645 Sequence 72645, A
9	146	39.9	376 15	US-10-282-122A-75772 Sequence 75772, A
10	144	39.3	428 15	US-10-282-122A-55748 Sequence 55748, A
11	142.5	38.9	372 15	US-10-282-122A-68109 Sequence 68109, A
12	141.5	38.7	323 15	US-10-282-122A-59321 Sequence 59321, A
13	141.5	38.7	347 14	US-10-127-032-120 Sequence 120, App

14	141.5	38.7	347	15	US-10-282-122A-66237	Sequence 66237, A
15	141	38.5	372	10	US-09-820-843A-8	Sequence 8, Appli
16	141	38.5	372	16	US-10-467-421-16	Sequence 16, Appli
17	140	38.3	389	15	US-10-282-122A-67145	Sequence 67145, A
18	132	36.1	507	16	US-10-726-692-50	Sequence 50, Appli
19	132	36.1	717	15	US-10-726-692-36	Sequence 36, Appli
20	131.5	35.9	336	15	US-10-282-122A-69962	Sequence 69962, A
21	131	35.8	45	9	US-09-816-989A-2	Sequence 2, Appli
22	130.5	35.7	388	15	US-10-282-122A-78190	Sequence 78190, A
23	129.5	35.4	214	14	US-10-229-567-27	Sequence 27, Appli
24	129.5	35.4	214	15	US-10-282-122A-62547	Sequence 62547, A
25	129.5	35.4	214	15	US-10-282-122A-64817	Sequence 64817, A
26	124.5	34.0	452	14	US-10-184-832-5	Sequence 5, Appli
27	124	33.9	212	15	US-10-282-122A-61735	Sequence 61735, A
28	123	33.6	433	15	US-10-390-472-2	Sequence 2, Appli
29	122.5	33.5	329	15	US-10-282-122A-45264	Sequence 45264, A
30	123	33.6	448	15	US-10-282-122A-67699	Sequence 67699, A
31	121.5	33.2	239	15	US-10-282-122A-60543	Sequence 60543, A
32	121	33.1	223	13	US-10-051-643-201	Sequence 201, App
33	121	33.1	223	14	US-10-205-979-52	Sequence 52, Appli
34	121	33.1	376	14	US-10-156-761-9889	Sequence 9889, Ap
35	120	32.8	240	17	US-10-739-930-9712	Sequence 9712, Ap
36	120	32.8	272	17	US-10-739-930-10710	Sequence 10710, A
37	120	32.8	568	17	US-10-435-115-197135	Sequence 197135, A
38	120	32.8	636	15	US-10-425-114-37076	Sequence 37076, A
39	120	32.8	926	16	US-10-437-963-193381	Sequence 193381, A
40	117	32.0	217	14	US-10-156-761-10221	Sequence 10221, A
41	117	32.0	279	14	US-10-181-071-7	Sequence 7, Appli
42	117	32.0	827	16	US-10-437-963-152005	Sequence 152005, A
43	116.5	31.8	885	17	US-10-425-115-198446	Sequence 198446, A
44	114	31.1	127	17	US-10-425-115-237481	Sequence 237481, A
45	112.5	30.7	146	15	US-10-296-115-1023	Sequence 1023, Ap

ALIGNMENTS

RESULT 1

US-09-816-989A-5
; Sequence 5, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKER
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-5

Query Match	100.0%;	Score 366;	DB 9;	Length 77;
Best Local Similarity	100.0%;	Pred. No. 9e+25;	Mismatches 0;	Indels 0;
Matches	77;	Conservative 0;	0;	Gaps 0;
QY	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKYAKAAEKKEYYAAAEAK	60	
Db	1	AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKAKAEAKKYAKAAEKKEYYAAAEAK	60	
QY	61	YKAEAKAAKAEAYEA 77		

```
Db          61 YKAEAAKAAKEAAVEA 77

RESULT 2
US-09-816-989A-6
; Sequence 6, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-6

Query Match          96.0%; Score 351.5; DB 9; Length 86;
Best Local Similarity 89.5%; Pred. No. 1.8e-23;
Matches 77; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY          1 AKKYAKKEKAYAKAEKAAKAAKAYKAAEAKKKAKAEAKKYAKAAKAEKKYAAAEAK 60
Db          1 AKKYAKKEKAYAKAEKAAKAAKAYKAAEAKKKAKAEAKKYAKAAKAEKKYAAAEAK 60
QY          61 -----YKAEAAKAAKEAAVEA 77
Db          61 YKAEAAKAYKAEAAKAAKEAAVEA 86

RESULT 3
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7

Query Match          79.0%; Score 289; DB 9; Length 109;
Best Local Similarity 67.0%; Pred. No. 5.8e-18;
Matches 73; Conservative 1; Mismatches 3; Indels 32; Gaps 4;

QY          1 AKKYAKK-EKAYAKA-----EKAAKKAEAKAYKAAEAKKKA----- 36
```

```
Db          1 AKKYAKKAEKAYAKKAAKAEKKAYAKKEAKAYKAAEAKKKAKAEAKKYAKAEAAKAKEA 60
QY          37 -KAEAKKYAKAAKAEKKEYAAAEAK-----YKAEAAKAAKAEAAVEA 77
Db          61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAKAEAAVEA 109

RESULT 4
US-09-816-989A-4
; Sequence 4, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-4

Query Match          77.2%; Score 282.5; DB 9; Length 66;
Best Local Similarity 84.4%; Pred. No. 1.3e-17;
Matches 65; Conservative 0; Mismatches 1; Indels 11; Gaps 2;

QY          1 AKKYAKKEKAYAKAEKAAKAEKAYKAAEAKKKAKAEAKKYAKAAKAEKKYAAAEAK 60
Db          1 AKKYAKKEKAYAK-----AKKAEAK-----AAKKAKEAKKYAKAAKAEKKYAAAEAK 49
QY          61 YKAEAAKAAKAEAAVEA 77
Db          50 YKAEAAKAAKAEAAVEA 66

RESULT 5
US-09-816-989A-3
; Sequence 3, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-3
```

[illegible]


```
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      39.3%; Score 144; DB 15; Length 428;
Best Local Similarity 59.0%; Pred. No. 8.3e-05;
Matches 49; Conservative 11; Mismatches 11; Indels 12; Gaps 6;

QY      1 AKYA--KKEKAYAKAEKAA-----KKAERAKYK--AAEAKKAKAEAKKYAKAAAEKK 52
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      167 AKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKA 225
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY      53 EYAAAEAKYKAAEA--KAAAEKAA 74
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      226 --AAAEAKKAAAEAKKAAEA 246
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 11
US-10-282-122A-68109
; Sequence 68109, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55748
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55748

Query Match      39.3%; Score 144; DB 15; Length 428;
Best Local Similarity 59.0%; Pred. No. 8.3e-05;
Matches 49; Conservative 11; Mismatches 11; Indels 12; Gaps 6;

QY      1 AKYA--KKEKAYAKAEKAA-----KKAERAKYK--AAEAKKAKAEAKKYAKAAAEKK 52
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      167 AKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKAADAQKKAEEAEAKKA 225
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY      53 EYAAAEAKYKAAEA--KAAAEKAA 74
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      226 --AAAEAKKAAAEAKKAAEA 246
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 12
US-10-282-122A-59321
; Sequence 59321, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match      38.9%; Score 142.5; DB 15; Length 372;
Best Local Similarity 46.9%; Pred. No. 9.7e-05;
Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY      1 AKKYAKKEKA-----YAKKAEEKAAKKA---EAKYKAAEAKKKAKAEAKKYA--- 44
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      148 AKKAAEKQADIAKKKAEEAKKAAEEAKKAAEEAKKAAEEAKKAAEEAKKAAEEAKKAAEEAKKAAED 207
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY      45 ---KKAKEKKEYYAAAEAKYK--AEAAKAAAEKAAEAYEA 77
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      208 AKKAAAEADAKKAAEEAKKAAADAAQKKAQAQAARKA 245
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

RESULT 12
US-10-282-122A-59321
; Sequence 59321, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match      38.9%; Score 142.5; DB 15; Length 372;
Best Local Similarity 46.9%; Pred. No. 9.7e-05;
Matches 46; Conservative 10; Mismatches 21; Indels 21; Gaps 4;

QY      1 AKKYAKKEKA-----YAKKAEEKAAKKA---EAKYKAAEAKKKAKAEAKKYA--- 44
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      148 AKKAAEKQADIAKKKAEEAKKAAEEAKKAAEEAKKAAEEAKKAAEEAKKAAEEAKKAAED 207
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||

QY      45 ---KKAKEKKEYYAAAEAKYK--AEAAKAAAEKAAEAYEA 77
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
DB      208 AKKAAAEADAKKAAEEAKKAAADAAQKKAQAQAARKA 245
      ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : ||| : : : |||
```

```
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59321
; LENGTH: 323
; TYPE: PRP
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59321

Query Match          38.7%; Score 141.5; DB 15; Length 323;
Best Local Similarity 56.3%; Pred. No. 0.0001;
Matches 40; Conservative 9; Mismatches 19; Indels 3; Gaps 2;

QY 8 EKAYAKAEKAAKAEKAYK-AAEAKKKAKAKAKYAKAAKAEKKEYAAAAEAKYAEAA 66
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 55 EQAAKAAADAKKQAEAAAKAAAEAKKQAEAAK--AAAEAKKAEAAAKKQAEAE 112
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
QY 67 KAAAEAAAYEA 77
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 113 KKAQQAARQA 123

RESULT 13
US-10-127-032-120
; Sequence 120, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; FILE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 120
; LENGTH: 347
; TYPE: PRP
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-120

Query Match          38.7%; Score 141.5; DB 14; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00011;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYAKAEKAAKAAKAAK-----YKAAE-A 32
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 120 ARKAEAKAEKAAKAEKAAKAEKAAEKKQADIAKKAEDAEKKAEDAKKKAEDA 179
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 33 KKAKEAEKAYAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 74
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 180 KKAEEAEAKKA-AAEAAKKAAVEAAKKAKEKKAKEKKAKEKKAKEKKAKEKKA 220
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

RESULT 14
US-10-282-122A-66237
; Sequence 66237, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
```

```
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Cart, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66237
; LENGTH: 347
; TYPE: PRP
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66237

Query Match          38.7%; Score 141.5; DB 15; Length 347;
Best Local Similarity 44.1%; Pred. No. 0.00011;
Matches 45; Conservative 9; Mismatches 19; Indels 29; Gaps 3;

QY 1 AKKYAKKEKAYAKAEKAAKAAKAAK-----YKAAE-A 32
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 120 ARKAEAKAEKAAKAEKAAKAEKAAEKKQADIAKKAEDAEKKAEDAKKKAEDA 179
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 33 KKAKEAEKAYAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 74
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 180 KKAEEAEAKKA-AAEAAKKAAVEAAKKAKEKKAKEKKAKEKKAKEKKAKEKKA 220
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

RESULT 15
US-09-820-843A-8
; Sequence 8, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 372
; TYPE: PRP
; ORGANISM: H. influenzae
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Outer membrane integrity protein (tolA)
```


This Page Blank (usp10)

Result No.	Query			ID	Description	
	Score	Match	Length			
1	366	100.0	77	4	US-09-405-743A-5	Sequence 5, Appli
2	351.5	96.0	86	4	US-09-405-743A-6	Sequence 6, Appli
3	289	79.0	109	4	US-09-405-743A-7	Sequence 7, Appli
4	282.5	77.2	66	4	US-09-405-743A-4	Sequence 4, Appli
5	245.5	67.1	56	4	US-09-405-743A-3	Sequence 3, Appli
6	144.5	39.5	100	2	US-08-460-890A-64	Sequence 64, Appl
7	144.5	39.5	100	3	US-08-167-641C-64	Sequence 64, Appl
8	144.5	39.5	100	3	US-08-460-971A-64	Sequence 64, Appl
9	144.5	39.5	100	3	US-08-462-040-64	Sequence 64, Appl
10	141.5	38.7	407	4	US-09-252-991A-29581	Sequence 29581, A
11	141.5	38.7	469	4	US-09-489-039A-13565	Sequence 13565, A
12	134	36.6	643	2	US-08-216-894-8	Sequence 8, Appli
13	134	36.6	643	3	US-09-115-746-8	Sequence 8, Appli
14	131	35.8	45	4	US-09-405-743A-2	Sequence 2, Appli
15	130	35.5	472	2	US-08-216-894-10	Sequence 10, Appl
16	130	35.5	472	3	US-09-115-746-10	Sequence 10, Appl
17	130	35.5	564	2	US-08-216-894-2	Sequence 2, Appli
18	130	35.5	564	3	US-09-115-746-2	Sequence 2, Appli
19	129.5	35.4	214	3	US-09-041-889-27	Sequence 27, Appl
20	129.5	35.4	214	4	US-09-417-264-27	Sequence 27, Appl
21	129.5	35.4	361	4	US-09-343-681A-5390	Sequence 5390, Ap
22	128.5	35.1	214	4	US-09-328-352-5169	Sequence 5169, Ap
23	125.5	34.3	468	4	US-09-328-352-6321	Sequence 6321, Ap
24	123	33.6	433	1	US-08-346-843-2	Sequence 2, Appli
25	123	33.6	433	2	US-08-293-284A-2	Sequence 2, Appli
26	123	33.6	433	4	US-08-898-300-2	Sequence 2, Appli
27	121	33.1	423	3	US-09-095-855-201	Sequence 201, App

RESULT 7
US-08-167-641C-64
; Sequence 64, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gotchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993

RESULT 8
US-08-460-971A-64
; Sequence 64, Application US/08460971A
; Patent No. 6150168
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,971A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389

APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13565
LENGTH: 469
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13565
Query Match 38.7%; Score 141.5; DB 4; Length 469;
Best Local Similarity 56.3%; Pred. No. 5.1e-06;
Matches 40; Conservative 9; Mismatches 19; Indels 3; Gaps 2;
QY 8 EKAYAKKAAKAKABAKAYK-AAEAKKKAKABAKKAYAKAAKAEKKEYAAAAEKYKAAEA 66
DB 205 EQAAKAAADAKQAAAAKAAKAAKQAAEAAK--AAAEAKKAAEAAAKKAAQQAEE 262
QY 67 KAAKEAAYEA 77
DB 263 KKAQQAQAQA 273
RESULT 12
US-08-216-894-8
Sequence 8, Application US/08216894
Patent No. 5876734
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,894
FILING DATE:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-216-894-8
Query Match 36.6%; Score 134; DB 2; Length 643;
Best Local Similarity 52.0%; Pred. No. 3.6e-05;
Matches 39; Conservative 9; Mismatches 23; Indels 4; Gaps 2;
QY 1 AKKYAKKEKAYAKAAKAAKAAKAYKAAEAKKKAAEAKKAAEAKKYAKAAK-AEKKEYAAAAEA 59
DB 490 ATKVAEAEK--QKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEA 546
QY 60 KYKAAEAKAAKAAEA 74
DB 547 TKVAEAEKQKAAEA 561
RESULT 14
US-09-405-743A-2
Sequence 2, Application US/09405743A
Patent No. 6514938
GENERAL INFORMATION:
APPLICANT: Yeda Research and Development Co., Ltd.
TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
FILE REFERENCE: 60807-A
CURRENT APPLICATION NUMBER: US/09/405,743A

DB 490 ATKVAEAEK--QKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEA 546
QY 60 KYKAAEAKAAKAAEA 74
DB 547 TKVAEAEKQKAAEA 561
RESULT 13
US-09-115-746-8
Sequence 8, Application US/09115746
Patent No. 6228601
GENERAL INFORMATION:
APPLICANT: Kirchhoff, Louis V.
APPLICANT: Otsu, Keiko
TITLE OF INVENTION: POLYPEPTIDES FOR DIAGNOSING INFECTION
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/115,746
FILING DATE:
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,894
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 85326/102/DRLO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-115-746-8
Query Match 36.6%; Score 134; DB 3; Length 643;
Best Local Similarity 52.0%; Pred. No. 3.6e-05;
Matches 39; Conservative 9; Mismatches 23; Indels 4; Gaps 2;
QY 1 AKKYAKKEKAYAKAAKAAKAAKAYKAAEAKKKAAEAKKAAEAKKYAKAAK-AEKKEYAAAAEA 59
DB 490 ATKVAEAEK--QKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEATKVAEAEKQKAAEA 546
QY 60 KYKAAEAKAAKAAEA 74
DB 547 TKVAEAEKQKAAEA 561
RESULT 14
US-09-405-743A-2
Sequence 2, Application US/09405743A
Patent No. 6514938
GENERAL INFORMATION:
APPLICANT: Yeda Research and Development Co., Ltd.
TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
FILE REFERENCE: 60807-A
CURRENT APPLICATION NUMBER: US/09/405,743A

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	164	40.1	421	2	JV0057	tolA protein - Esc	
2	157.5	38.5	376	2	AG0592	tolA protein [impo	
3	156.5	38.3	1701	2	T09127	probable erythrocy	
4	155	37.9	394	2	F90725	membrane spanning	
5	155	37.9	394	2	G85576	membrane spanning	
6	154	37.7	347	2	E83525	TolA protein PA097	
7	151	36.9	210	2	A28550	histone H1 - sea u	
8	149	36.4	231	2	S59589	histone H1 - Chlam	
9	147.5	36.1	388	2	AC0138	TolA colicin impo	
10	147	35.9	214	2	G70673	probable hupB - My	
11	144	35.2	211	2	A28100	histone H1-beta, e	
12	142.5	34.8	344	2	S34153	mst101-1 protein -	
13	142	34.7	372	2	G64064	outer membrane int	
14	140.5	34.4	328	2	A44993	cytosolic repetiti	
15	140.5	34.4	1390	2	S51364	sperm tail-specifi	
16	138	33.7	217	2	A26721	histone H1-gamma,	
17	137.5	33.6	1128	2	T30296	R27-2 protein - Tr	
18	135	33.0	219	2	E60110	repetitive protein	
19	134	32.8	206	2	S09388	histone H1 - sea u	
20	133.5	32.6	384	2	B43592	outer membrane pro	
21	133.5	32.6	433	2	S25194	zuotin - yeast [sa	
22	132.5	32.4	206	1	HSTR1P	histone H1 - rainb	
23	132	32.3	248	1	HSUR1P	histone H1, gonada	
24	131	32.0	208	2	T23778	histone H1.1 - Cae	
25	130.5	31.9	220	2	A28456	histone H1.10 - ch	
26	130.5	31.9	243	2	AE1689	hypothetical prote	
27	129.5	31.7	924	2	T06636	hypothetical prote	
28	129.5	31.7	1403	2	T11583	probable translati	
29	129	31.5	218	2	A23055	histone H1.01 - ch	

Search completed: December 14, 2004, 05:52:15
Job time : 20.7764 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 96.5232 Seconds
(without alignments)
512.646 Million cell updates/sec

Title: US-10-792-311-6

Perfect score: 409
Sequence: 1 AKTYAKKEKAYAKKAEKAAK.....KKAYKAEAAKAAKAEAAVEA 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :      UniProt_02:*
1:  uniprot_sprot:*
2:  uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	165.5	40.5	40.5	407	2	Q8ZQT6	Q8zqt6 salmonella
2	165	40.3	37.3	1701	2	Q7RC08	Q7rc08 plasmodium
3	164	40.1	42.1	1	1	TOLA ECOLI	P19934 escherichia
4	163	39.9	37.2	2	2	Q9WXX1	Q9wwx1 pseudomonas
5	163	39.9	37.2	2	2	Q88N16	Q88n16 pseudomonas
6	162	39.6	41.3	2	2	Q7C2Q4	Q7c2q4 shigella fl
7	162	39.6	41.3	2	2	Q83SA1	Q83sa1 shigella fl
8	160.5	39.2	39.5	2	2	Q6D7F3	Q6d7f3 erwinia car
9	160	39.1	42.1	2	2	Q8FUT1	Q8fjt1 escherichia
10	157.5	38.5	37.6	2	2	Q8ZSC1	Q8zsc1 salmonella
11	156.5	38.3	165.2	2	2	Q7KPEZ1	Q7kpe21 plasmodium
12	156.5	38.3	165.2	2	2	AAQ73456	AAq73456 plasmodium
13	156.5	38.3	1680	2	2	Q7KPEY9	Q7kpy9 plasmodium
14	156.5	38.3	1680	2	2	AAQ73455	AAq73455 plasmodium
15	156.5	38.3	1680	2	2	AAQ73457	AAq73457 plasmodium
16	156.5	38.3	1701	2	2	O61164	O61164 plasmodium
17	155	37.9	37.9	2	2	Q7WFN5	Q7wfn5 bordetella
18	155	37.9	39.4	2	2	Q7AG18	Q7agi8 escherichia
19	155	37.9	39.4	2	2	Q8X965	Q8x965 escherichia
20	154	37.7	34.7	1	1	TOLA PSEAE	P50600 pseudomonas
21	154	37.7	37.3	2	2	Q7W477	Q7w477 bordetella
22	154	37.7	39.5	2	2	Q937K4	Q937k4 erwinia chr
23	153	37.4	40.1	2	2	Q74W64	Q74w64 yersinia pe
24	153	37.4	40.1	2	2	AAS61283	AAs61283 yersinia
25	152.5	37.3	71.3	2	2	Q6CDX0	Q6cdx0 yarrowia li
26	151	36.9	21.0	1	1	H1 LYPTI	P06144 lytechinus
27	151	36.9	1341	2	2	Q98KG7	Q98kg7 rhizobium l
28	150.5	36.8	198	2	2	Q6MIU4	Q6mtu4 bdellovibri
29	150.5	36.8	198	2	2	CAE80819	CAe80819 bdellovib
30	148.5	36.3	232	2	2	Q39576	Q39576 chlamydomo
31	148	36.2	389	2	2	Q9CM70	Q9cm70 pasteurella

RESULT 1

RESOLUT 1
Q8ZQT6
ID Q8ZOT6
PRELIMINARY:
PRT: 407 AA.

AC	Q8ZQTS		
AD	01-MAR-2002	(TRENburel. 20, Created)	
DT	01-MAR-2002	(TRENburel. 20, last sequence update)	
DT	01-MAR-2004	(TRENburel. 26, last annotation update)	
DE	Tol protein, membrane spanning protein.		
DE	Name=colA; OrderedLocusNames=STM0747;		
GN	Salmonella typhimurium.		
OS	Salmonella typhimurium.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriaceae; Salmonella.		
OC	Enterobacteriaceae; Salmonella.		

OX	NCBI_TaxID=602;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=L22;	
RX	MEDLINE=21534948; Pubmed=11677609;	
RA	McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.	
RA	Courtney L., Powellik S., Ali J., Dante M., Du F., Hou S., Layman D.	
RA	Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,	
RA	Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,	
RA	Waterston R., Wilson R.K.;	
RT	Complete genome sequence of <i>Salmonella enterica</i> serovar Typhimurium	

Query Match	40.5%	Score 165.5;	DB 2;	Length 407;
Best Local Similarity	51.9%	Pred. No. 0.00018;		
Matches	55;	Conservative	14;	Mismatches 16;
				Indels 21;
				Gaps 6;

[illegible]

QY 52 KEYAAAEAKYKAE-----AKKAYKAEAAKAA--KEAAYEA 86
: |||: || ||| ||| : ||||| ||| : || ||
183 AK-AAADAKKKKAEAAKAAAEAKKKKAEAAKAAAEAKKKKADAE 221

RESULT 2

Q7RC08	PRELIMINARY;	PRT; 1701 AA.
ID	Q7RC08	
AC	Q7RC08;	
DT	01-MAR-2004	(TRENBLrel. 26, Created)
DT	01-MAR-2004	(TRENBLrel. 26, Last sequence update)
DT	01-MAR-2004	(TRENBLrel. 26, Last annotation update)
DE	Erythrocyte binding protein.	
GN	Name=PY05977;	

```

OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
NCBI_TaxID=73239;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carleton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteau M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoaihi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy_binding.
DR Pfam; PF05424; Duffy_binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 40.3%; Score 165; DB 2; Length 1701;
Best Local Similarity 56.1%; Pred. No. 0.00064;
Matches 55; Conservative 18; Mismatches 18; Indels 18; Gaps 6;

QY 1 AKKYAKKEKA-YAKKAEKAKKAEKAYKAEKAKKAKA-----EAKKYAKAAK---AEK 51
Db 1270 AKKAEKKKAAKAAKAEKKKAEKAEKAKKAEKAEKKKAEKAEKAEKAEKAEKAEK 1328
QY 52 KEVAAA-----EAKYKAEKAAK-----YKAEKAAKAAKE 81
Db 1329 KKAEEAKKAEKKKAEKAEKAEKAEKKKAEKAEKKKAEKAEKAEKAEKAEKAEK 1366

RESULT 3
TOLA_ECOLI STANDARD; PRT; 421 AA.
AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Tola protein.
GN Name:tola; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levengood S.K., Webster R.E.;
RT "Nucleotide sequences of the tola and tolB genes and localization of
RT their products, components of a multistep translocation system in
RT Escherichia coli."
RL J. Bacteriol. 171:6600-6609(1989).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

```

RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:1137-1155(1996).
[4]
RN DOMAINS.
RP MEDLINE=91296736; PubMed=2068069;
RX Levengood S.K., Beyer W.F. Jr., Webster R.E.;
RT "Tola: a membrane protein involved in colicin uptake contains an
RT extended helical region."
RL Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
[5]
RN INTERACTION WITH PORINS.
RP MEDLINE=97133271; PubMed=8978668;
RX Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "Tola central domain interacts with Escherichia coli porins."
RL EMBO J. 15:6408-6415(1996).
[6]
RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RP MEDLINE=99332679; PubMed=10404600;
RX Lubkowski J., Hennecke F., Pluckthun A., Wlodawer A.;
RA "Filamentous phage infection: crystal structure of g3p in complex with
RA its coreceptor, the C-terminal domain of Tola."
RL Structure 7:711-722(1999).
CC -!- FUNCTION: Involved in the tonB-independent uptake of group A
CC colicins (colicins A, E1, E2, E3, and K). Necessary for the
CC colicins to reach their respective targets after initial binding
CC to the bacteria. Also involved in the translocation of
CC bacteriophage DNA.
CC -!- SUBUNIT: Interacts, via domain II, with porins ompC, phoE and
CC lamb.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M28232; AAA24683.1; -.
CC EMBL; A5000177; AAC73833.1; -.
CC EMBL; D50713; BAA35405.1; -.
CC PIR; JY0057; JY0057.
CC PDB; ITOL; X-ray; A=--
CC EchoBASE; EB1000; -.
CC EcoGene; EG11007; tola.
CC InterPro; IPR010528; Tola.
CC Pfam; PF06519; Tola; 1.
KM 3D-structure; Bacteriocin transport; Complete proteome;
KW Inner membrane; Protein transport; Repeat; Transmembrane; Transport.
FT DOMAIN 1 13 Cytoplasmic (Potential).
FT TRANSMEM 14 34 Potential.
FT DOMAIN 35 421 Periplasmic (Potential).
FT DOMAIN 48 310 DOMAIN II (ALPHA-HELICAL).
FT DOMAIN 311 421 DOMAIN III (FUNCTIONAL).
FT DOMAIN 224 252 13 tandem repeats of [EDA]-K(1,2)-A(2,4).
FT REPEAT 224 229 1.
FT REPEAT 230 234 2.

```

```

FT REPEAT 235 240 3.
FT REPEAT 241 245 4.
FT REPEAT 246 250 5.
FT REPEAT 251 255 6.
FT REPEAT 256 260 7.
FT REPEAT 261 266 8.
FT REPEAT 267 271 9.
FT REPEAT 272 277 10.
FT REPEAT 278 282 11.
FT REPEAT 283 287 12.
FT REPEAT 288 292 13.
FT DISULFID 363 368
FT HELIX 335 349
FT TURN 350 351
FT TURN 353 354
FT TURN 355 358
FT TURN 359 360
FT TURN 363 369
FT TURN 371 372
FT STRAND 375 383
FT HELIX 385 397
FT HELIX 406 412
FT TURN 413 414
FT STRAND 416 421
SQ SEQUENCE 421 AA; 43156 MW; 8B2F52B4B97C655E CRC64;

Query Match 40.1%; Score 164; DB 1; Length 421;
Best Local Similarity 59.3%; Pred. No. 0.00024;
Matches 51; Conservative 7; Mismatches 18; Indels 10; Gaps 4;

QY 7 KEKAYAKAEKAAKAEKAYKAAEAK-KKAAEAKKAYKAA-----KAEKKEVAAAEAA 59
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 148 KAEADAKAEAEAKKAAADAKKKAEEAKAAAEAAQKKAEEAAALKKKAAEAAEAAAEAA 207
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 60 KYKA--EAAKAYKAEAAKAAAEKAA 83
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 208 RKAAATEAAEKA-KAEAEKAAAEKAA 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 4
Q9WWX1 PRELIMINARY; PRT; 372 AA.
AC Q9WWX1
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Tola protein.
CN Name=tola;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96198174; PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
lipoprotein (PAL) is involved in maintenance of the integrity of the
cell envelope."
RL J. Bacteriol. 178:1699-1706 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX MEDLINE=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida."
RL J. Bacteriol. 178:5836-5840 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez I.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

```

```

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; F:nucleosome assembly; IEA.
DR GO; GO:0015031; F:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; Tola.
DR InterPro; IPR006260; TonB_C.
DR Pfam; PF06519; Tola; 1.
DR PRINTS; PRO0624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match 39.9%; Score 163; DB 2; Length 372;
Best Local Similarity 53.1%; Pred. No. 0.00025;
Matches 52; Conservative 12; Mismatches 22; Indels 12; Gaps 5;

QY 1 AKKYAKKEKA--YAKKAE-KAAKAEAKAYKAA--EAKKKAAEAKKYA-----KAKKA 49
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 148 AKKAAEKQADIAKKKAEDEAKKAAEEAKKAAAEAKKAAEDAKKAAEEAKKKAED 207
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 50 EKKEVAAAEKVK-AEAAKAYKAAEAKAAAEAAEAA 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 208 AKKAAAEADAKKAAEEAKKAAADAQKKKAQEAARKA 245
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 5
Q88NI6 PRELIMINARY; PRT; 372 AA.
AC Q88NI6
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Biopolymer transport protein Tola.
GN Name=tola; OrderedLocusNames=PP1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Pouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Maazee A., Uteerback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmli B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016778; AAN66845.1; -.
DR HSSP; P50600; 1LR0.
DR TIGR; PP1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.

```

```

DR GO: GO:0006334; P:nucleosome assembly; IEA.
DR GO: GO:0015031; P:protein transport; IEA.
DR InterPro: IPR005819; Histone_H5.
DR InterPro: IPR010528; Tola.
DR InterPro: IPR006260; TonB_C.
DR Pfam: PF06519; Tola; 1.
DR PRINTS: PR00624; HISTONH5.
DR TIGRPFAMs: TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785ECC3C0BC CRC64;

Query Match
Best Local Similarity 53.1%; Pred. No. 0.00025;
Matches 52; Conservative 12; Mismatches 22; Indels 12; Gaps 5;

QY 1 AKKYAKKEKA--YAKKAE-KAAKKAAYKAA--EAKKAKAEAKYA-----KAAKA 49
Db 148 AKKAAEKQQAADIAKKAADEAKKAAEEAKKAAAEAKKAAEDAKKAAAEAKKAAED 207
QY 50 EKKEVAAAEAKYK-AEAAKAYKAAEAKKAAAEAKKAAAEAKKAAAEAKKAAAE 86
Db 208 AKKAAAEADAKKAAAEAKKAAADAKKAAAEAKKAAAEAKKAAAEAKKAAAE 245

RESULT 6
Q7C2Q4 PRELIMINARY; PRT; 413 AA.
AC Q7C2Q4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Membrane spanning protein.
GN Name=tola; OrderedLocusNames=S0571;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016979; AAP16075.1; -.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
SQ SEQUENCE 413 AA; 42355 MW; 93E10F2C5DE60DE8 CRC64;

Query Match
Best Local Similarity 39.6%; Score 162; DB 2; Length 413;
Matches 52; Conservative 7; Mismatches 23; Indels 10; Gaps 4;

QY 1 AKKYAKKEKAYAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 53
Db 134 AEVAAAKAAADAKAAEEAKKAAADAKKAAEAAKAAAEAKKAAEAAKAAAEAAK 193
QY 54 YAAAEAKYKA--EAAKKAYKAAEAKKAAAEAKKAAAEAKKAAAEAKKAAAEAK 83
Db 194 AAAAEARKKAATEAAEKA-KAAEAKKAAAEKA 224

RESULT 7
Q83SA1 PRELIMINARY; PRT; 413 AA.
AC Q83SA1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

QY 1 AKKYAKKEKAYAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 53
Db 134 AEVAAAKAAADAKAAEEAKKAAADAKKAAEAAKAAAEAKKAAEAAKAAAEAAK 193
QY 54 YAAAEAKYKA--EAAKKAYKAAEAKKAAAEAKKAAAEAKKAAAEAKKAAAEAK 83
Db 194 AAAAEARKKAATEAAEKA-KAAEAKKAAAEKA 224

RESULT 8
Q6D7F3 PRELIMINARY; PRT; 395 AA.
AC Q6D7F3;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tola protein.
GN Name=tola; ORFNames=ECAL372;
OS Erwinia carotovora subsp. atroseptica SCR11043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCR11043;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrall B.G., Parkhill J., Toth I.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG74282.1; -.
SQ SEQUENCE 395 AA; 41698 MW; 2DB1DAA64E984199 CRC64;

Query Match
Best Local Similarity 39.2%; Score 160.5; DB 2; Length 395;
Matches 51; Conservative 11; Mismatches 18; Indels 11; Gaps 5;

QY 5 AKKEKAYAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 60
Db 134 AQOAKEQQKQAAEAAKAAKAAEAAKAAEAAKAAEAAKAAEAAKAAEAAKAAEAAK 192
QY 61 YK-AEAAK-----KAYKAAEAKAAAEAAEAAEAAEAAEAAEAAEAAEAAEAA 86

```

```

DE Membrane spanning protein, required for outer membrane integrity.
DE Name=tola; OrderedLocusNames=SF0558;
GN Shigella flexneri
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015086; AAN42202.1; -.
DR HSSP; P19934; 1TOL.
DR InterPro: IPR010528; Tola.
DR Pfam: PF06519; Tola; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 42355 MW; 93E10F2C5DE60DE8 CRC64;

Query Match
Best Local Similarity 39.6%; Score 162; DB 2; Length 413;
Matches 52; Conservative 7; Mismatches 23; Indels 10; Gaps 4;

QY 1 AKKYAKKEKAYAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 53
Db 134 AEVAAAKAAADAKAAEEAKKAAADAKKAAEAAKAAAEAKKAAEAAKAAAEAAK 193
QY 54 YAAAEAKYKA--EAAKKAYKAAEAKKAAAEAKKAAAEAKKAAAEAKKAAAEAK 83
Db 194 AAAAEARKKAATEAAEKA-KAAEAKKAAAEKA 224

RESULT 8
Q6D7F3 PRELIMINARY; PRT; 395 AA.
AC Q6D7F3;
DT 01-OCT-2004 (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Tola protein.
GN Name=tola; ORFNames=ECAL372;
OS Erwinia carotovora subsp. atroseptica SCR11043.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=218491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SCR11043;
RA Bell K.S., Sebaihia M., Pritchard L., Holden M., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagsels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Barrall B.G., Parkhill J., Toth I.K.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX950851; CAG74282.1; -.
SQ SEQUENCE 395 AA; 41698 MW; 2DB1DAA64E984199 CRC64;

Query Match
Best Local Similarity 39.2%; Score 160.5; DB 2; Length 395;
Matches 51; Conservative 11; Mismatches 18; Indels 11; Gaps 5;

QY 5 AKKEKAYAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 60
Db 134 AQOAKEQQKQAAEAAKAAKAAEAAKAAEAAKAAEAAKAAEAAKAAEAAKAAEAAK 192
QY 61 YK-AEAAK-----KAYKAAEAKAAAEAAEAAEAAEAAEAAEAAEAAEAAEAA 86

```



```
Query Match      38.3%; Score 156.5; DB 2; Length 1680;
Best Local Similarity 53.5%; Pred. No. 0.0026;
Matches 53; Conservative 9; Mismatches 18; Indels 19; Gaps 6;
```

```
Qy    1 AKKYAKTEKAYA----KKAERKAAKKAEAKYKAAEAKTKAKA-----BAKKYAKAAK--- 48
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   1204 AAKKAEERKKAEAVKKAEEAKKCAEA-AKKAEERKKCAEAAKKALRRKKKSBAACKALE 1362
```

```
Qy    49 AEKKEYAA--AEAKYKAEAKKA-----YKAEAAKAAAAKE 81
      ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db   1263 RKKKAEAAKKAEEKKKCAEAAKKAAEEKKKCAEAAKKAEEEE 1301
```

Search completed: December 14, 2004, 06:10:19
Job time : 97.5232 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 95.0717 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-6
Perfect score: 409
Sequence: 1 AKYAKKEKAYAKAEKAAK.....KKAYKAEKAAKAEAYEA 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	100.0	86	3	AAY82576 Copolymer
2	351.5	85.9	77	3	AAY82575 Copolymer
3	312.5	76.4	109	3	AAY82577 Copolymer
4	268	65.5	66	3	AAY82574 Copolymer
5	233	57.0	56	3	AAY82573 Copolymer
6	165.5	40.5	407	6	ABU47123 Protein e
7	184	40.1	387	6	ABU44721 Protein e
8	164	40.1	421	6	ABU28559 Protein e
9	163	39.9	154	2	AAR06445 Recombina
10	163	39.9	372	6	ABU40185 Protein e
11	161	39.4	428	6	ABU27824 Protein e
12	157.5	38.5	323	6	ABU11397 Protein e
13	157.5	38.5	376	6	ABU47848 Protein e
14	157.5	38.5	469	7	ABO67048 Klebsiell
15	157	38.4	106	2	AAR06446 Recombina
16	156	38.1	100	3	AAY98499 Peptide #
17	156	38.1	100	3	AAY59044 Amino aci
18	156	38.1	100	4	AAB45852 Nucleic a
19	156	38.1	100	4	AUU04289 Poly-L-lys
20	154	37.7	347	6	ABU38313 Protein e
21	154	37.7	347	6	ABJ18771 Pseudomon
22	154	37.7	407	7	ABO80835 Pseudomon
23	148	36.2	389	6	ABU39221 Protein e
24	147.5	36.1	388	6	ABU50266 Protein e
25	147	35.9	214	2	AAY34055 M. tuberc

26	147	35.9	214	2	AAY57353	Aay57353 M. tuberc
27	147	35.9	214	6	ABU34623	Abu34623 Protein e
28	147	35.9	214	6	ABU36893	Abu36893 Protein e
29	147	35.9	357	6	ABM67869	Abm67869 Phototrab
30	146.5	35.8	205	3	AAB20575	Aab20575 Mycobacte
31	142	34.7	372	5	ABG80418	Abg80418 Haemophil
32	142	34.7	372	7	ABO23507	AbO23507 Haemophil
33	141.5	34.6	336	6	ABU42038	Abu42038 Protein e
34	141.5	34.6	361	7	ADF05105	Adf05105 Bacterial
35	140.5	34.4	214	6	ADA33882	Ada33882 Acinetoba
36	140	34.2	468	6	ADA35034	Ada35034 Acinetoba
37	138.5	33.9	80	5	ABG71044	Abg71044 Tumour ne
38	137.5	33.6	448	6	ABU17340	Abu17340 Protein e
39	137.5	33.6	472	2	AAR84569	Aar84569 Trypanoso
40	137.5	33.6	507	8	ADQ36692	Adq36692 Trypanoso
41	137.5	33.6	564	2	AAR84565	Aar84565 Trypanoso
42	137.5	33.6	643	2	AAR84568	Aar84568 Trypanoso
43	137.5	33.6	717	8	ADQ36678	Adq36678 Trypanoso
44	135	33.0	140	2	AAR25206	Aar25206 Synthetic
45	133.5	32.6	433	2	AAW30256	Aaw30256 Zootin. 3

ALIGNMENTS

RESULT 1
AAY82576
ID AAY82576 standard; peptide; 86 AA.
XX
AC AAY82576;
XX
DT 28-JUL-2000 (first entry)
XX
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.
XX
KW Copolymer; molecular weight marker; TV-marker; immune disease;
KW Glairamer acetate; autoimmune disease; antiarthritic; neuroprotective;
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
KW antidabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
KW pemphigus vulgaris; systemic lupus erythematosus.
XX
OS Unidentified.
XX
FN WO200018794-A1.
XX
PD 06-APR-2000.
XX
PF 24-SEP-1999; 99WO-US022402.
XX
PR 25-SEP-1998; 98US-0101693P.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
PA (TEVA-) TEVA PHARM USA INC.
XX
PI Gad A, Lis D;
XX
DR WPI; 2000-317499/27.
XX
PT Copolymer 1 related polypeptides used as molecular weight markers for
PT glairamer acetate and for treatment and prevention of immune diseases.
XX
PS Claim 10; Page 14; 72pp; English.
XX
CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular
CC weight TV-marker polypeptides from the present invention. The present
CC invention describes polypeptides (I) for determining the molecular weight
CC of a copolymer (CP), which has an identified molecular weight and an
CC amino acid composition corresponding to the copolymer. The polypeptides

Job time : 96.0717 secs

```
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67145
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-67145

Query Match      36.2%; Score 148; DB 15; Length 389;
Best Local Similarity 47.2%; Pred.No. 4.7e-05;
Matches 50; Conservative 11; Mismatches 23; Indels 22; Gaps 5;

QY      2 KKYAKKEKAYAKAEKAAK-KAEAKAYKAA-----EAKKKAKAEAKKYAKAEKAE-- 50
Db      142 EKQKQAEAKAKQLAEAAKLEAEAKRLAALAKQAEAEKAKAEAEKAKAEKAEKAE 201

QY      51 -----KKEYAAAEAKYKAEAAK-----KAYKAEAAKAAKAEAAKAE 86
Db      202 AKAKVEKAKAEAEAKVKAEKAEAEAKA-KAEKAKAEAEKAEKAEKAE 246
```

Search completed: December 14, 2004, 06:59:42
Job time : 214.367 secs

This Page Blank (uspio)

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	409	100.0	86	4	US-09-405-743A-6	Sequence 6, Appli	
2	351.5	85.9	77	4	US-09-405-743A-5	Sequence 5, Appli	
3	312.5	76.4	109	4	US-09-405-743A-7	Sequence 7, Appli	
4	268	65.0	66	4	US-09-405-743A-4	Sequence 4, Appli	
5	233	57.5	56	4	US-09-405-743A-3	Sequence 3, Appli	
6	157.5	38.5	469	4	US-09-489-039A-13565	Sequence 13565, A	
7	156	38.1	100	2	US-08-460-890A-64	Sequence 64, Appl	
8	156	38.1	100	3	US-08-167-641C-64	Sequence 64, Appl	
9	156	38.1	100	3	US-08-460-971A-64	Sequence 64, Appl	
10	156	38.1	100	3	US-08-462-040-64	Sequence 64, Appl	
11	154	37.7	407	4	US-09-252-991A-29581	Sequence 29581, A	
12	147	35.9	214	3	US-09-041-889-27	Sequence 27, Appl	
13	147	35.9	214	4	US-09-417-264-27	Sequence 27, Appl	
14	141.5	34.6	361	4	US-09-843-681A-5390	Sequence 5390, Ap	
15	140.5	34.4	214	4	US-09-328-352-5169	Sequence 5169, Ap	
16	140	34.2	468	4	US-09-328-353-6321	Sequence 6321, Ap	
17	137.5	33.6	472	2	US-08-216-894-10	Sequence 10, Appl	
18	137.5	33.6	472	3	US-09-115-746-10	Sequence 10, Appl	
19	137.5	33.6	584	2	US-08-216-894-2	Sequence 2, Appli	
20	137.5	33.6	564	3	US-09-115-746-2	Sequence 2, Appli	
21	137.5	33.6	643	2	US-08-216-894-8	Sequence 8, Appli	
22	137.5	33.6	643	3	US-09-115-746-8	Sequence 8, Appli	
23	133.5	32.6	433	1	US-08-346-849-2	Sequence 2, Appli	
24	133.5	32.6	433	2	US-08-293-284A-2	Sequence 2, Appli	
25	133.5	32.6	433	4	US-08-898-300-2	Sequence 2, Appli	
26	129.5	31.7	223	3	US-09-095-855-201	Sequence 201, App	
27	129.5	31.7	223	4	US-09-205-426-201	Sequence 201, App	


```

205 EQAAKAAADAKKQAEAAAKAAAEAKKQAEAAK--AAAEQKKAEEAAAKKQAEAE 262
Db

67 KKAYKAEAAKAAAEKAAAYE 85
QY
|||:|:||||:|
263 KKAQQAANKQAEEAKAAAE 281
Db

```

```

RESULT 7
US-08-460-890A-64
; Sequence 64, Application US/08460890A
; Patent No. 5994109
; GENERAL INFORMATION:
; APPLICANT: Woo, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen A.
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Fastseq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,890A
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/167,641
; FILING DATE: December 14, 1993
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 212/066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: "Lys Ala" in positions 3 to 100 may
; OTHER INFORMATION: present or absent.
US-08-460-890A-64

```

[illegible]

Db 62 AKAKAKAKAKAKAKAKAKAKAKA 86

RESULT 8
US-08-167-641C-64
; Sequence 64, Application US/08167641C
; Patent No. 6033884
; GENERAL INFORMATION:
; APPLICANT: WOO, Savio L.C.
; APPLICANT: Smith, Louis C.
; APPLICANT: Cristiano, Richard J.
; APPLICANT: Gottchalk, Stephen
; TITLE OF INVENTION: NUCLEIC ACID TRANSPORTER SYSTEMS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,641C
; FILING DATE: December 14, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/855,389
; FILING DATE: March 20, 1992
; APPLICATION NUMBER: PCT/US93/02725
; FILING DATE: March 19, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 205/012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; OTHER INFORMATION: "Lys Ala" in positions 3 to 100 may
; ; OTHER INFORMATION: present or absent.
US-08-167-641C-64

```

Query Match      38.1%; Score 156; DB 3; Length 100;
Best Local Similarity 55.3%; Pred. No. 7.7e-08;
Matches 47; Conservative 6; Mismatches 30; Indels 2; Gaps 2;

Qy 1 AKKYAK-KEKAYAKAEKAAKAEKAYAKAEAKGKAKAEAKYAKA-AKAEKKEYAAAAE 58
    || || || || || || || || || || || || || || || || || || || || ||
Db 2 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAK 61
    || || || || || || || || || || || || || || || || || || || || ||

Qy 59 AKYKAEAAKAYKAEAAKAAAEAA 83
    || || || || || || || || || || || || || || || || || || || || ||
Db 62 AKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKAKA 86
    || || || || || || || || || || || || || || || || || || || || ||

```

RESULT 9

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 37.7%; Score 154; DB 4; Length 407;
Best Local Similarity 47.6%; Pred. No. 5e-07;
Matches 49; Conservative 11; Mismatches 21; Indels 22; Gaps 4;
QY 1 AKKYAKKERAKYAKAEKAAKAEKAYKAA-----EAKKAKAEAKKYA---- 44
Db 180 ARKAEQAQKAAEAKKAEKAEKAAEKKQADIAKRAEDEAKKAAEDAKKAAEDA 239
QY 45 --KAAYAEKYEAAAEAKYKA--EAKKAYKAEAAKAAKAA 83
Db 240 KKAAAEAKKAAAEAAKAAKAAVEAAK--KAAAEAAKAAKAA 280

RESULT 12
US-09-041-889-27
; Sequence 27, Application US/09041889
; Patent No. 6033864
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,889
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/837,058
; FILING DATE: 11-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-041-889-27

Query Match 35.9%; Score 147; DB 3; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.1e-06;
Matches 47; Conservative 3; Mismatches 34; Indels 10; Gaps 3;
QY 1 AKKYAKKERAKYAKAEKAAKAEKAEK-----AYKAAEAKKKAKAEKAYKAAKAEKKEY 54
Db 111 AKKVAK--KAPAKKATKAAKKAATKAPAKKAATKAPAKKAATKAPAKKAATKAPAKKV 168
QY 55 AAAEAK--YKAEAAKAYKAEAAKAAKAAEAAVEA 86
Db 169 TKAVKTKAVKASVRKAATKAPAKKAAAKRPATKA 202
RESULT 13
US-09-417-264-27
; Sequence 27, Application US/09417264
; Patent No. 6537768
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/417,264
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/041,889
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-417-264-27

Query Match 35.9%; Score 147; DB 4; Length 214;
Best Local Similarity 50.0%; Pred. No. 1.1e-06;
Matches 47; Conservative 3; Mismatches 34; Indels 10; Gaps 3;
QY 1 AKKYAKKERAKYAKAEKAAKAEKAEK-----AYKAAEAKKKAKAEKAYKAAKAEKKEY 54
Db 111 AKKVAK--KAPAKKATKAAKKAATKAPAKKAATKAPAKKAATKAPAKKAATKAPAKKV 168
QY 55 AAAEAK--YKAEAAKAYKAEAAKAAKAAEAAVEA 86
Db 169 TKAVKTKAVKASVRKAATKAPAKKAAAKRPATKA 202
RESULT 14
US-09-543-681A-5390

```
; Sequence 5390, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5390
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match      34.6%; Score 141.5; DB 4; Length 361;
Best Local Similarity 52.1%; Pred. No. 6.3e-06;
Matches 50; Conservative 12; Mismatches 19; Indels 15; Gaps 6;

QY      5 AKKEKAYAKAEAKAEAK-----AYKAEAKKAKAEAKKYAKAA--KAEK-KEYAA 56
Db      138 AAKAKBEQQAEEAAQAERDRILKEQADAKAKAEAEAKQAELAAKQAEAKAE 197

QY      57 AEAKYKAEAKKAYKAEAE---AKAAAK--EAAAYEA 86
Db      198 AEAKAKAEADAKA-KAEADAKAKAEAKAEAKAAAEA 232

RESULT 15
US-09-328-352-5169
; Sequence 5169, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5169
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5169

Query Match      34.4%; Score 140.5; DB 4; Length 214;
Best Local Similarity 51.0%; Pred. No. 4.5e-06;
Matches 50; Conservative 10; Mismatches 23; Indels 15; Gaps 6;

QY      1 AKKYAKKEK--AYAKAEAKAEAKAEAKYKAAAEAKKYKAEAKKYAKAAK--AEKKEYA- 55
Db      29 AKRLAEADKKAEAEAKQAEADKKA-AEAKQAEADKKA-AEAKQAEADKKAEAEAKQAE 86

QY      56 ----AAEAKYKAEAKKA-----YKAEAAKAAAEAAVE 85
Db      87 ADKKAEEAKQAEADKKAEAEAKKAEAEKAEAEKARE 124

Search completed: December 14, 2004, 05:50:18
Job time : 24.2737 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 25.0654 Seconds
(without alignments)
418.411 Million cell updates/sec

Title: US-10-792-311-7
Perfect score: 519
Sequence: 1 AKKYAKAEKAYAKAKAAK.....AKAYKAEAKAAKAEAYEA 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 segs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.5	36.5	421	JV0057	tola protein - Esc
2	181.5	35.0	394	F90725	membrane spanning
3	181.5	35.0	394	G85576	membrane spanning
4	181	34.9	210	A23550	histone H1 - sea u
5	179	34.5	347	E83525	Tola protein PA097
6	178	34.3	376	AG0592	tola protein [impo
7	178	34.3	388	AC0138	Tola colicin impor
8	175.5	33.8	344	S34153	mst101-1 protein -
9	172.5	33.2	211	A28100	histone H1-beta, e
10	171.5	33.0	1390	S51364	sperm tail-specifi
11	170.5	32.9	248	HSUR1P	histone H1, gonada
12	170.5	32.9	1701	T09127	probable erythrocy
13	165	31.8	220	A28456	histone H1.10 - ch
14	164.5	31.7	231	S59589	histone H1 - Chlam
15	164	31.6	265	S19113	cgr-4 protein - C
16	164	31.6	311	T17698	hypothetical prote
17	163.5	31.5	384	B43592	outer membrane pro
18	163	31.4	182	S61926	histone H1 homolog
19	163	31.4	225	B28456	histone H1.11L - c
20	162.5	31.3	1128	T30296	R27-2 protein - Tr
21	162	31.2	356	A82152	tola protein VC193
22	161.5	31.1	328	A44993	cytosolic repetiti
23	161	31.0	219	E60110	repetitive protein
24	160.5	30.9	206	HSR11R	histone H1 - rainb
25	160.5	30.9	226	S51660	histone H1-5 [vali
26	159.5	30.7	219	C28456	histone H1.11R - c
27	158	30.4	219	HSRUIB	histone H1-4 [vali
28	157	30.3	771	A33430	h-caldesmon - chic
29	156.5	30.2	218	A23055	histone H1.01 - ch

ALIGNMENTS

RESULT 1

JV0057
tola protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: JV0057; B64810
R:Levengood, S.K.; Webster, R.E.
J. Bacteriol. 171, 6600-6609, 1989
A>Title: Nucleotide sequences of the tola and tolB genes and localization of their products
A:Reference number: JV0057; MUID:90078104; PMID:2687247
A:Accession: JV0057
A:Molecule type: DNA
A:Residues: 1-421 <LEV>
A:Cross-references: UNIPROT:P19934; GB:M28232; NID:G148018; PIDN:AAA24683.1; PID:G148019
A:Experimental source: strain JM105
A>Note: The authors translated the initiation codon GTG for residue 1 as Val
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Coi
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A>Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64810
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-421 <BLAT>
A:Cross-references: GB:AE000177; GB:U00096; NID:G1786955; PIDN:AAC73833.1; PID:G1786960;
A:Experimental source: strain K-12, substrain MG1655
C:Comment: tola and tolB proteins are necessary for colicins E2, E3, A, and K to reach th
C:Genetics:
A:Gene: tola
A:Map position: 17 min
A:Start codon: GTG
C:Keywords: nucleotide binding; P-loop; transmembrane protein
F:14-34/Domain: transmembrane #status predicted <MSS>
F:78-301/Domain: helical #status predicted <HSR>
F:355-362/Region: nucleotide-binding motif A (P-loop)

Query Match 36.5%; Score 189.5; DB 2; Length 421;
Best Local Similarity 53.0%; Pred. No. 1.7e-05;
Matches 61; Conservative 11; Mismatches 32; Indels 11; Gaps 5;

QY 1 AKKYAKAE-----KAYAKAKAAEKKAYAKAEAKKAKAEAKKAYAK-
120 AEEAAKQAELEKQKQAEAAAKAAADAKAKAEADAKA--AEEAAKKAADAKKAEAAK 177
Db
QY 56 AKKEAY-KAEAKKYAKAKAEKKEVYAAAEAKK----AEAAKAYKAEAAKAAAEKAA 106
178 AAEEAQKKAEEAAALKKKAEAAAEAAAEAAEAAKKAATEAAEKAKAEAKKAAAEKA 232
Db

RESULT 2

F90725

Matches	54;	Conservative	11;	Mismatches	37;	Indels	15;	Gaps	4;
---------	-----	--------------	-----	------------	-----	--------	-----	------	----

```
QY 1 AKKYAKKAEKAY-----AKKAKAAKKEKKYAKKAEKAYKAAE-----AKKKAKAKKAY 49
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 1067 AKKLKEAAEKQCEERAKKEKEAAEKQCEERAKKLKEAAEKQCEERAKKEKEAAEKKR 1126
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 50 AKEAAKAKKAEKAYKAEAKKAYKAAKAEKKEVYAAAEAKKAEAAKAYKAAEAAKAAKAEAA 106
: ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 1127 CEEAAKREKEA--AEKKCAEAAKKE--ATEKQKCAEAAKKEKEAAEKKKCAEAA 1179
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||

RESULT 11
HSUR1P
histone H1, gonadal - sea urchin (Parechinus angulosus)
C:Species: Parechinus angulosus (angulate urchin)
C>Date: 31-Mar-1980 #sequence_revision 31-Mar-1980 #text_change 09-Jul-2004
C:Accession: A91090; A91091; A02586
R:Strickland, W.N.; Strickland, M.; de Groot, P.C.; von Holt, C.; Wittmann-Liebold, B.
Eur. J. Biochem. 104, 559-566, 1980
A:Title: The primary structure of histone H1 from sperm of the sea urchin Parechinus angulosus
A:Reference number: A91090; MUID:80156831; PMID:6767609
A:Contents: sequence of residues 1-84
A:Accession: A91090
A:Molecule type: protein
A:Residues: 1-248 <STR>
A:Cross-references: UNIPROT:P02256
R:Strickland, W.N.; Strickland, M.; Brandt, W.F.; von Holt, C.; Lehmann, A.; Wittmann-Liebold, B.
Eur. J. Biochem. 104, 567-578, 1980
A:Title: The primary structure of histone H1 from sperm of the sea urchin Parechinus angulosus
A:Reference number: A91091; MUID:80156832; PMID:7363905
A:Accession: A91091
A:Molecule type: protein
A:Residues: 80-248 <ST2>
A:Note: 144-Arg was also found
C:Superfamily: histone H1
C:Keywords: DNA binding; nucleosome; sperm

Query Match 32.9%; Score 170.5; DB 1; Length 248;
Best Local Similarity 46.0%; Pred. No. 0.00018;
Matches 52; Conservative 9; Mismatches 41; Indels 11; Gaps 2;

QY 2 KKYAKKAEKAYKAAKAEKKKAYKAEKAYKAAEAKKAEKAYKAAEAKKAYKAAEAKKAYKAAEAKKAY 61
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 120 KKAATSAAKKAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 179
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 62 KA--EAKYAK-----AAKAEKVEYAAAEAKKAEAAKAYKAAEAAKAAK 103
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 180 KAAKAKKPAKSPKKAKKPAKSPKKKAKRSPKKAKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 232
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||

RESULT 12
T09127
probable erythrocyte-binding protein MAEBL - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09127
R:Kapke, S.H.I.; Noe, A.R.; Fraser, T.S.; Blair, P.L.; Adams, J.H.
Proc. Natl. Acad. Sci. U.S.A. 95, 1230-1235, 1998
A:Title: A family of chimeric erythrocyte binding proteins of malaria parasites.
A:Reference number: Z16577; MUID:98115903; PMID:9448314
A:Accession: T09127
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1701 <KAP>
A:Cross-references: UNIPROT:O61164; EMBL:AF031886; NID:g2947227; PID:g2947228
A:Experimental source: subspecies yoelii; strain YM
C:Genetics:
A:Gene: mael
A:Introns: 62/1; 1648/1; 1674/2; 1697/1
C:Keywords: alternative splicing; cell binding; erythrocyte invasion

Query Match 32.9%; Score 170.5; DB 2; Length 1701;
Best Local Similarity 53.5%; Pred. No. 0.00076;
```

```
Matches 54; Conservative 11; Mismatches 31; Indels 5; Gaps 5;
QY 5 AKKAEKAYKAAKAEKKAYAKKAEKAYKAAEAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 63
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 1205 AKKAEKE--RKGAEVK--KAEAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 1261
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 64 EAKYAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 104
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 1262 ERKKAEAAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 1301
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||

RESULT 13
A28456
histone H1.10 - chicken
C:Species: Gallus gallus (chicken)
C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 04-Mar-2000
C:Accession: A28456
R:Colles, L.S.; Robins, A.J.; Madley, L.K.; Wells, J.R.E.
J. Biol. Chem. 262, 9656-9663, 1987
A:Title: Characterization of the chicken histone H1 gene complement. Generation of a complementary DNA library
A:Reference number: A92655; MUID:87250632; PMID:3597432
A:Accession: A28456
A:Molecule type: DNA
A:Residues: 1-220 <COL>
A:Cross-references: GB:M17018; NID:g211834; PIDN:AAA48788.1; PID:g211835
C:Superfamily: histone H1
C:Keywords: acetylated amino end; chromosomal protein; DNA binding; nucleosome; nucleus
P;2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

Query Match 31.8%; Score 165; DB 2; Length 220;
Best Local Similarity 46.4%; Pred. No. 0.00038;
Matches 52; Conservative 11; Mismatches 41; Indels 8; Gaps 3;

QY 1 AKKYAKKAEKAYKAAKAEKKKAYKAEKAEKAYKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 53
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 109 SKKPGVEKKEKPRKRTPAKPKKPAKPKKPAKPKKPAKPKKPAKPKKPAKPKKPAKPKKPAKPKKPA 168
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
QY 54 AKAEKAEKAYKAEKAEKAYKAAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 104
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||
Db 169 AKSPKAYKAAKPKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 220
||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : ||| ||| : |||

RESULT 14
S59589
histone H1 - Chlamydomonas reinhardtii
C:Species: Chlamydomonas reinhardtii
C>Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59589; S62122
R:Fabry, S.; Mueller, K.; Lindauer, A.; Park, P.B.; Cornelius, T.; Schmitt, R.
Curr. Genet. 28, 333-345, 1995
A:Title: The organization structure and regulatory elements of Chlamydomonas histone gene
A:Reference number: S59581; MUID:96120862; PMID:8590479
A:Accession: S59589
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-231 <PAB>
A:Cross-references: UNIPROT:Q39576; EMBL:U16726
A:Note: the authors did not translate the codon for residue 1
R:Fabry, S.; Ford, C.; Lindauer, A.; Mueller, K.; Cornelius, T.; Schmitt, R.
submitted to the EMBL Data Library, October 1994
A:Description: The organization, structure and controlling elements of Chlamydomonas histone gene
A:Reference number: S62122
A:Accession: S62122
A:Molecule type: DNA
A:Residues: 1-173, P'174-231 <PAW>
A:Cross-references: EMBL:U16726; NID:g571479; PIDN:AAA98452.1; PID:g571480
C:Genetics:
A:Introns: 62/3; 101/3
C:Superfamily: histone H1
C:Keywords: chromosomal protein; DNA binding; nucleosome

Query Match 31.7%; Score 164.5; DB 2; Length 231;
Best Local Similarity 45.9%; Pred. No. 0.00042;
```


Search completed: December 14, 2004, 05:52:15
Job time : 25.0654 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 122.338 Seconds
(without alignments)
512.646 Million cell updates/sec

Title: US-10-792-311-7
Perfect score: 519
Sequence: 1 AKYAKAEKAYAKAKAAK.....AKYAKAEAKAAKAEAYEA 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_02: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	202	38.9	395	2	Q937K4	Q937k4 erwinia chr
2	193.5	37.3	1701	2	Q7RC08	Q7rc08 plasmodium
3	190.5	36.7	372	2	Q9WWX1	Q9wwx1 pseudomonas
4	190.5	36.7	372	2	Q8BN16	Q8bn16 pseudomonas
5	190.5	36.7	1817	2	Q7KS08	Q7ks08 plasmodium
6	190.5	36.7	1817	2	AAQ73467	AAq73467 plasmodium
7	190.5	36.7	1830	2	Q7K5Q7	Q7k5q7 plasmodium
8	190.5	36.7	1830	2	AAQ73466	AAq73466 plasmodium
9	190.5	36.7	1866	2	Q8T5C8	Q8t5c8 plasmodium
10	189.5	36.5	421	1	TOLA ECOLI	P19934 escherichia
11	188.5	36.3	421	2	Q8EJ71	O8fjt1 escherichia
12	186.5	35.9	401	2	Q74W64	Q74w64 yersinia pe
13	186.5	35.9	401	2	AA561283	AA561283 yersinia
14	184.5	35.5	407	2	Q8ZQT6	Q8zqt6 salmonella
15	181.5	35.0	394	2	Q8AGI8	Q8agi8 escherichia
16	181.5	35.0	394	2	Q8X965	Q8x965 escherichia
17	181	34.9	210	1	H1 LYTP1	P06144 lytechinus
18	179	34.5	347	1	TOLA_PSEAE	P50600 pseudomonas
19	179	34.5	395	2	Q6D7F3	Q6d7f3 erwinia car
20	178	34.3	376	2	Q8Z8C1	Q8z8c1 salmonella
21	178	34.3	388	2	Q8ZGZ2	Q8zg22 yersinia pe
22	178	34.3	393	2	Q8CZ28	Q8cz28 yersinia pe
23	177.5	34.2	420	2	Q73D27	Q73d27 bacillus ce
24	177.5	34.2	420	2	AA539818	AA539818 bacillus
25	177	34.1	713	2	Q6CDX0	Q6cdx0 yarrowia li
26	176.5	34.0	200	2	Q8XVW7	Q8xvn7 ralistonia s
27	176.5	34.0	413	2	Q7C2Q4	Q7c2q4 shigella fl
28	176.5	34.0	413	2	Q83SA1	Q83sa1 shigella fl
29	176	33.9	177	2	Q6SG84	Q6sg84 uncultured
30	176	33.9	177	2	AA37978	AA37978 unculture
31	176	33.9	1268	2	Q6CGN4	Q6cgn4 yarrowia li

32	175.5	33.8	344	1	MST1_DROHY	Q08695 drosophila
33	173.5	33.4	389	2	Q9CM70	Q9cm70 pasteurella
34	173.5	33.4	524	2	Q6HHE7	Q6hhe7 bacillus th
35	171.5	33.0	1391	1	MST2_DROHY	Q08696 drosophila
36	171	32.9	436	2	Q6F986	Q6f986 acinetobact
37	170.5	32.9	248	1	H1 PARAN	P02256 parechinus
38	170.5	32.9	1652	2	O7KP21	O7kp21 plasmodium
39	170.5	32.9	1652	2	AAQ73456	AAq73456 plasmodium
40	170.5	32.9	1680	2	Q7KPY9	Q7kpy9 plasmodium
41	170.5	32.9	1680	2	AAQ73455	AAq73455 plasmodium
42	170.5	32.9	1680	2	AAQ73457	AAq73457 plasmodium
43	170.5	32.9	1701	2	O61164	O61164 plasmodium
44	170	32.8	232	2	Q39576	Q39576 chlamydomon
45	169.5	32.7	211	1	H1B_STRPU	P15869 strongyloce

ALIGNMENTS

RESULT 1

Q937K4 PRELIMINARY; PRT; 395 AA.

AC Q937K4; (TREMBlrel. 19, Created)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

DE Tola protein.

DE Tola protein.

GN Name=tola;

OS Erwinia chrysanthemi.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Pectobacterium.

OX NCBI_TaxID=556;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=3937;

RA Ray M.C., Vianney A., Cotte-pattat N., Lazzaroni J.;

RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ297885; CAC82708.1; -.

DR HSSP; P19934; ITOL.

DR InterPro; IPR010528; Tola.

DR Pfam; PF06519; Tola; 1.

SQ SEQUENCE 395 AA; 41601 MW; 3C0C1DC12E181013 CRC64;

Query Match 38.9%; Score 202; DB 2; Length 395;

Best Local Similarity 47.3%; Pred. No. 6.5e-06;

Matches 62; Conservative 15; Mismatches 32; Indels 22; Gaps 4;

QY 1 AKYAKAEKAYAK-----KAKAEKKAYAKAEKAYAKAEAKKKAK-----43

Db 138 AKEQQKQAEVAAKAKAEAEQQAQAAADAKKQAEVEVKQAADAKKKAEAEAKTKAAAA 197

QY 44 AEAKKYAKAEAKKKEA---YKAEAKYAKAAKAEKKEVYAAAEAKKAEAAKAYK--AEAA 98

Db 198 AEAKKAEAEKAKAAADAKKQAEAEAKKAEAKAEKAAADAAKAEAAAKKAAADDK 257

QY 99 KAAAEKAEAYEA 109

Db 258 KKAATAAKQA 268

RESULT 2

Q7RC08 PRELIMINARY; PRT; 1701 AA.

AC Q7RC08; (TREMBlrel. 26, Created)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)

DE Erythrocyte binding protein.

GN Name=PY05977;

OS Plasmodium yoelii yoelii.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=73239;

RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegh M., Shoaibi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B., O.R.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii.";
RL Nature 419:512-519(2002)
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01001968; EAA18109.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 1701 AA; 197910 MW; 03BE665BAE45C669 CRC64;

Query Match 37.3%; Score 193.5; DB 2; Length 1701;
Best Local Similarity 50.4%; Pred. No. 8.1e-05;
Matches 60; Conservative 12; Mismatches 28; Indels 19; Gaps 5;

QY 2 KTKAKAEKAYAKAKAEKKAY-AKKEAKYAKAEAKKAKA-----EAKYAKAEAK 55
Db 1238 KKAEAAKALERKKAEAKKALERKKAEAKKAEKKAEKKAEKKAEKKAEAKK 1297
QY 56 AKKEAYAKAEKKYAKAEKKYAA-----BAKKAEEK-----AYKAEAAKAAKE 104
Db 1298 ABEKKKAEKAEKAEKKAEKKAEKKAEKKAEKKAEKKAEKKAEKKAEKKAE 1353

RESULT 3
Q9WXX1 PRELIMINARY; PRT; 372 AA.
AC Q9WXX1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TolA protein.
GN Name=tolA;
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX PubMed=8626299;
RA Rodriguez-Herva J.J., Ramos-Gonzalez M.I., Ramos J.;
RT "The Pseudomonas putida peptidoglycan-associated outer membrane
RT lipoprotein (PAL) is involved in maintenance of the integrity of the
RT cell envelope.";
RL J. Bacteriol. 178:1699-1706(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RX PubMed=96422022; PubMed=8824639;
RA Rodriguez-Herva J.J., Ramos J.;
RT "Characterization of an OprL null mutant of Pseudomonas putida.";
RL J. Bacteriol. 178:5836-5840(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Ramos-Gonzalez I.;

```

```

RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=mt-2;
RA Rodriguez-Herva J.J.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; X74218; CAB50780.1; -.
DR HSSP; P50600; 1LR0.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006334; P:nucleosome assembly; IEA.
DR GO; GO:0015031; P:protein transport; IEA.
DR InterPro; IPR005819; Histone_H5.
DR InterPro; IPR010528; TolA.
DR Pfam; PF06519; TolA; 1.
DR PRINTS; PR00624; HISTONEH5.
DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785EC3C0BC CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 372;
Best Local Similarity 50.4%; Pred. No. 3.5e-05;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

QY 5 AKKAEKAYAKKAAK-----EKKAYAKKAEKAYKAE-----AKKAEKAYKAE 54
Db 118 AKKAEDA-AKAEAAKAAKAEKAAKAEKAAKAEKAAKAEKAAKAEKAAKAE 176
QY 55 K-AKKEAYAKKAYAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 109
Db 177 KKAEEAKKAAEDAKKKAEEAKKAAEDAKKKAEEAKKAAEDAKKKAEEAKKAA 233

RESULT 4
Q88NI6 PRELIMINARY; PRT; 372 AA.
AC Q88NI6
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Biopolymer transport protein TolA.
GN Name=tolA; OrderedLocustNames=PP1221;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzez A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedter H., Lauber J., Stjepandic D., Hohseisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoef A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016778; AAN66845.1; -.
DR HSSP; P50600; 1LR0.
DR TIGR; PP1221; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0000786; C:nucleosome; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.

```

```
DR GO: GO:0008565; P:protein transporter activity; IEA.
DR GO: GO:0006334; P:nucleosome assembly; IEA.
DR GO: GO:0015031; P:protein transport; IEA.
DR InterPro: IPR005819; Histone_H5.
DR InterPro: IPR010528; Tola.
DR InterPro: IPR006260; TonB_C.
DR Pfam: PF06519; Tola; 1.
DR PRINTS: PR00624; HISTONEH5.
DR TIGRFAMs: TIGR01352; tonB_Cterm; 1.
KW Complete proteome.
SQ SEQUENCE 372 AA; 40133 MW; 87F49785BCC3C0BC CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 372;
Best Local Similarity 50.4%; Pred. No. 3.5e-05;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

Qy 5 AKKAEKAYAKKAAKAE-----KKYAKKEAKYKAAE-----AKKKAKEAKYKAAE 54
Db 118 AKKAEDA-AKAAEAKAAEAKAAEAKKADAEKKAQADIAKKKADEAKKAEAA 176
Qy 55 K--AKKEAYKAEAKYKAAKAEKKEYYAAAEAKKAAEAKKAAEAKKAAEAKKAAEAA 109
Db 177 KKAEEAEAKKAEDAKKKAEEAEAKKAAEDAKKKAEEAEAKKAAEAKKAAEAKKAAADA 233

RESULT 5
Q7K5Q8 PRELIMINARY; PRT; 1817 AA.
AC Q7K5Q8
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 3.
GN Name=maebl;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73467.1; -.
DR InterPro: IPR011046; WD40 like.
SQ SEQUENCE 1817 AA; 206998 MW; CE2C737CD537B553 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1817;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

Qy 5 AKKAEKAYAKKAAKAE-----KKYAKKEAKYKAAEAKKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAAEAKKAEAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1319
Qy 51 KEAAKAKKEAYKAEAKYKAAKAEKKEYYAAAEAKKAAEAKKAEAAK-----KAEAKKYA 106
Db 1320 EDARKA-EARKAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1376

RESULT 6
AAQ73467 PRELIMINARY; PRT; 1817 AA.
AC AAQ73467
DT 02-WAR-2004 (TRENBLrel. 27, Created)
DT 02-WAR-2004 (TRENBLrel. 27, Last sequence update)
DT 02-WAR-2004 (TRENBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 3.

GN MAEBL.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Salvador;
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73467.1; -.
SQ SEQUENCE 1817 AA; 206998 MW; CE2C737CD537B553 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1817;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

Qy 5 AKKAEKAYAKKAAKAE-----KKYAKKEAKYKAAEAKKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAAEAKKAEAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1319
Qy 51 KEAAKAKKEAYKAEAKYKAAKAEKKEYYAAAEAKKAAEAKKAEAAK-----KAEAKKYA 106
Db 1320 EDARKA-EARKAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1376

RESULT 7
Q7K5Q7 PRELIMINARY; PRT; 1830 AA.
AC Q7K5Q7
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN Name=maebl;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73466.1; -.
DR InterPro: IPR011046; WD40 like.
SQ SEQUENCE 1830 AA; 208650 MW; 099B2D035391E399 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1830;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

Qy 5 AKKAEKAYAKKAAKAE-----KKYAKKEAKYKAAEAKKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAAEAKKAEAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1319
Qy 51 KEAAKAKKEAYKAEAKYKAAKAEKKEYYAAAEAKKAAEAKKAEAAK-----KAEAKKYA 106
Db 1320 EDARKA-EARKAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAA 1376

RESULT 8
107 YEA 109
1377 RKA 1379
```

```
AAQ73466
ID AAQ73466 PRELIMINARY; PRT; 1830 AA.
AC AAQ73466;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-WAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-WAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN MAEBL.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Salvador;
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAQ73466.1; -.
SQ SEQUENCE 1830 AA; 208690 MW; 099B2D035391E399 CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1830;
Best Local Similarity 49.6%; Pred. No. 0.00013;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

QY 5 AKKAEKAYAKKAAKE-----KKAYAKKAYKAAEAKKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAEAKKAEARKAEARKAEARKAEARKAEARKAEARKAEARKAEARKA 1319
QY 51 KEAAKAKKAYKAAKAYKAAKAEKKEYAAAEAKKAEAAK-----AYKAEAAKAAK 106
Db 1320 EDARKA--BEARKAEAAKAEARKAEARKAEARKAEARKAEARKAEARKAEARKA--EAA 1376
QY 107 YEA 109
Db 1377 RKA 1379

RESULT 9
Q8T5C8 PRELIMINARY; PRT; 1866 AA.
ID Q8T5C8;
AC Q8T5C8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erythrocyte binding protein 1.
GN Name=maebi;
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5855;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
DR EMBL; AY042083; AAL10508.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffiy_binding.
DR Pfam; PF05424; Duffiy_binding; 1.
SQ SEQUENCE 1866 AA; 212420 MW; DC692D7CFAE7D93F CRC64;

Query Match 36.7%; Score 190.5; DB 2; Length 1866;
Best Local Similarity 49.6%; Pred. No. 0.00014;
Matches 61; Conservative 17; Mismatches 22; Indels 23; Gaps 6;

QY 5 AKKAEKAYAKKAAKE-----KKAYAKKAEKAYKAAEAKKKA-----KAEAKKYA 50
Db 1262 AKKAE--ARKAEAKKAEARKAEARKAEARKAEARKAEARKAEARKAEARKA 1319

AAQ73466
ID 51 KEAAKAKKAYKAAKAYKAAKAEKKEYAAAEAKKAEAAK-----AYKAEAAKAAK 106
Db 1320 EDARKA--BEARKAEAAKAEARKAEARKAEARKAEARKAEARKAEARKAEARKA--EAA 1376
QY 107 YEA 109
Db 1377 RKA 1379

RESULT 10
TOLA_ECOLI
ID TOLA_ECOLI STANDARD; PRT; 421 AA.
AC P19934;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Tola protein.
GN Name=tolA; Synonyms=cim, excC, lky; OrderedLocusNames=b0739;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / JM105;
RX MEDLINE=90078104; PubMed=2687247;
RA Levensgood S.K., Webster R.E.;
RT "Nucleotide sequences of the tolA and tolB genes and localization of
their products, components of a multistep translocation system in
Escherichia coli.";
RL J. Bacteriol. 171:6600-6609(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP DOMAINS.
RX MEDLINE=91296736; PubMed=2068069;
RA Levensgood S.K., Beyer W.F. Jr., Webster R.E.;
RT "TolA: a membrane protein involved in colicin uptake contains an
extended helical region.";
RN [5]
RP Proc. Natl. Acad. Sci. U.S.A. 88:5939-5943(1991).
RP INTERACTION WITH PORINS.
RX MEDLINE=97133271; PubMed=8978668;
RA Derouiche R., Gavioli M., Benedetti H., Prilipov A., Lazdunski C.,
RA Lloubes R.;
RT "TolA central domain interacts with Escherichia coli porins.";
RL EMBO J. 15:6408-6415(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 298-421.
RX MEDLINE=99332679; PubMed=10404600;
RA Lubkowski J., Henneke F., Plueckthun A., Wlodawer A.;
```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 120.498 Seconds
(without alignments)
324.499 Million cell updates/sec

Title: US-10-792-311-7
Perfect score: 519
Sequence: 1 AKYAKKAEKAYAKKAAK.....AKYAKAEAKAAKAAEAYEA 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	519	100.0	109	3	AAY82577 Copolymer
2	312.5	60.2	86	3	AAY82576 Copolymer
3	289	55.7	77	3	AAY82575 Copolymer
4	228.5	44.0	66	3	AAY82574 Copolymer
5	209	40.3	154	2	AAR06445 Recombina
6	196	37.8	469	7	ABO67048 Klebsiell
7	190.5	36.7	372	6	ABU40185 Protein e
8	190	36.6	428	6	ABU27824 Protein e
9	189.5	36.5	421	6	ABU28559 Protein e
10	187	36.0	106	2	AAR06446 Recombina
11	184.5	35.5	407	6	ABU47123 Protein e
12	181.5	35.0	323	6	ABU31397 Protein e
13	180.5	34.8	56	3	AAY82573 Copolymer
14	179.5	34.6	100	3	AAY98499 Peptide #
15	179.5	34.6	100	3	AAY59044 Amino aci
16	179.5	34.6	100	4	AAB45852 Nucleic a
17	179.5	34.6	100	4	AU04289 Poly-Lys-
18	179	34.5	347	6	ABU38313 Protein e
19	179	34.5	347	6	ABJ18771 Pseudomon
20	179	34.5	407	7	ABO80835 Pseudomon
21	178	34.3	223	2	ABY14928 Amino aci
22	178	34.3	223	6	ABP70903 Mycobacte
23	178	34.3	376	6	ABU47848 Protein e
24	178	34.3	388	6	ABU50266 Protein e
25	174	33.5	387	6	ABU44721 Protein e

26	173.5	33.4	389	6	ABU39221 Protein e
27	169	32.6	361	7	ADF05105 Bacterial
28	163.5	31.5	448	6	ABU17340 Protein e
29	163	31.4	336	6	ABU42038 Protein e
30	162.5	31.3	643	2	AAR84568 Trypanoso
31	162	31.2	356	6	ABU49418 Protein e
32	162	31.2	356	7	ABO23526 Vibrio ch
33	161	31.0	357	6	ABM67869 Photorhab
34	160.5	30.9	158	2	AAY34068 Histone H
35	160.5	30.9	158	2	AAY57366 Human his
36	160.5	30.9	226	2	AAY34060 Human his
37	160.5	30.9	226	2	AAY57358 Human his
38	160.5	30.9	226	8	ADQ09196 Human HLF
39	160.5	30.9	468	6	ADA35034 Acinetoba
40	160	30.8	214	6	ADA33882 Acinetoba
41	159.5	30.7	234	5	AAE13234 Human lin
42	159.5	30.7	234	5	AAU09944 Histone lin
43	159.5	30.7	234	7	ADD22892 Human his
44	158.5	30.5	472	2	AAR84569 Trypanoso
45	158	30.4	218	2	AAW29477 Human his

ALIGNMENTS

RESULT 1
AAY82577
ID AAY82577 standard; peptide; 109 AA.
XX
AC AAY82577;
XX
DT 28-JUL-2000 (first entry)
XX
DE Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:7.
XX
KW Copolymer; molecular weight marker; TV-marker; immune disease;
KW glatiramer acetate; autoimmune disease; antirheumatic; neuroprotective;
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
KW antianemic; immunosuppressive; demyelinating disease; osteoarthritis;
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
KW Crohn's disease; chronic immune thrombocytopaenia purpura; colitis;
KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
KW pemphigus vulgaris; systemic lupus erythematosus.
XX
OS Unidentified.

PN WO200018794-A1.
XX
PD 06-APR-2000.
XX
PF 24-SEP-1999; 99WO-US022402.
XX
PR 25-SEP-1998; 98US-0101693P.
XX
PA (VEDA) YEDA RES & DEV CO LTD.
PA (TEVA-) TEVA PHARM USA INC.
XX
PI Gad A, Lis D;
XX
WPI; 2000-317499/27.
XX
PT Copolymer 1 related polypeptides used as molecular weight markers for
PT glatiramer acetate and for treatment and prevention of immune diseases.
XX
PS Claim 10; Page 14; 72pp; English.

CC AAY82571 to AAY82577 represent specifically claimed copolymer molecular
CC weight TV-marker polypeptides from the present invention. The present
CC invention describes polypeptides (I) for determining the molecular weight
CC of a copolymer (CP), which has an identified molecular weight and an
CC amino acid composition corresponding to the copolymer. The polypeptides

CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC ophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 109 AA;
 Query Match 100.0%; Score 519; DB 3; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.9e-36;
 Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AKKYAKAEKAYAKKAAEKKEKAYAKKEAKKAYAKAEAKKAYAKAEAKKAAKEA 60
 DB 1 AKKYAKAEKAYAKKAAEKKEKAYAKKEAKKAYAKAEAKKAYAKAEAKKAAKEA 60
 QY 61 YKAEAKKYAKAAKAEKKEKAYAKAEAKKAEAKKAYAKAEAKKAAKAAEAYEA 109
 DB 61 YKAEAKKYAKAAKAEKKEKAYAKAEAKKAEAKKAYAKAEAKKAAKAAEAYEA 109
 RESULT 2
 AA82576
 ID AA82576 standard; peptide; 86 AA.
 AC AA82576;
 DT 28-JUL-2000 (first entry)
 XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:6.
 DE
 XX Copolymer; molecular weight marker; TV-marker; immune disease;
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 DE Crohn's disease; chronic immune thrombocytopenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.
 XX Unidentified.
 OS
 XX WO200018794-A1.
 PN
 XX 06-APR-2000.
 PD
 XX 24-SEP-1999; 99WO-US022402.
 PF
 XX 25-SEP-1998; 98US-0101693P.
 PR
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA (TEVA-) TEVA PHARM USA INC.
 XX
 XX Gad A, Lis D;
 PI
 XX WPI; 2000-317499/27.
 DR
 XX Copolymer 1 related polypeptides used as molecular weight markers for
 PT glatiramer acetate and prevention of immune diseases.

XX Claim 10; Page 14; 72pp; English.
 PS
 XX
 CC AA82571 to AA82577 represent specifically claimed copolymer molecular
 CC weight TV-marker polypeptides from the present invention. The present
 CC invention describes polypeptides (I) for determining the molecular weight
 CC of a copolymer (CP), which has an identified molecular weight and an
 CC amino acid composition corresponding to the copolymer. The polypeptides
 CC of the invention are used as molecular weight markers for glatiramer
 CC acetate related tetrapolymers. The polypeptides may also be used for
 CC treating and preventing immune diseases in a mammal. Autoimmune diseases
 CC which may be treated include either cell-mediated or antibody-mediated
 CC diseases. Such diseases include arthritic conditions, demyelinating
 CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
 CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
 CC ophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
 CC disease, chronic immune thrombocytopenia purpura, colitis, contact
 CC sensitivity disease, diabetes mellitus, Graves disease, Guillain-Barre's
 CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
 CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
 CC mediated diseases which can be treated include host-versus-graft disease,
 CC graft-versus-host disease, and delayed-type hypersensitivity. The
 CC polypeptides of the invention have defined molecular weights and physical
 CC properties which are analogous to glatiramer acetate molecules, which
 CC makes them ideal for use as molecular weight markers
 XX
 SQ Sequence 86 AA;
 Query Match 60.2%; Score 312.5; DB 3; Length 86;
 Best Local Similarity 72.1%; Pred. No. 5e-19;
 Matches 80; Conservative 1; Mismatches 3; Indels 27; Gaps 5;
 QY 1 AKKYAKAEKAYAKKAAEKKEKAYAKKEAKKAYAKAEAKKAEAKKAAKEA 60
 DB 1 AKKYAKK-ERAYAKKA-----EKAACKAEKAYAKAEAKKKA----- 36
 QY 61 YKAEAKKYAKAAKAEKKEKAYAKAEAKKAEAKKAEAKKAEAKKAAEAYEA 109
 DB 37 -KAEAKKYAKAAKAEKKEKAYAKAEAKKAEAKKAEAKKAEAKKAAEAYEA 86
 RESULT 3
 AA82575
 ID AA82575 standard; peptide; 77 AA.
 AC AA82575;
 DT 28-JUL-2000 (first entry)
 XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:5.
 DE
 XX Copolymer; molecular weight marker; TV-marker; immune disease;
 KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
 KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
 KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
 KW antianaemic; immunosuppressive; demyelinating disease; osteoarthritis;
 KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
 DE Crohn's disease; chronic immune thrombocytopenia purpura; colitis;
 KW diabetes mellitus; Graves disease; Guillain-Barre's syndrome; psoriasis;
 KW Hashimoto's disease; idiopathic myxoedema; myasthenia gravis;
 KW pemphigus vulgaris; systemic lupus erythematosus.
 XX Unidentified.
 OS
 XX WO200018794-A1.
 PN
 XX 06-APR-2000.
 PD
 XX 24-SEP-1999; 99WO-US022402.
 PF
 XX 25-SEP-1998; 98US-0101693P.
 PR
 XX (YEDA) YEDA RES & DEV CO LTD.
 PA (YEDA) YEDA RES & DEV CO LTD.

```
PA (TEVA-) TEVA PHARM USA INC.
XX
PI Gad A, Lis D;
XX
DR WPI; 2000-317499/27.
XX
DR Copolymer 1 related polypeptides used as molecular weight markers for
PT glatiramer acetate and for treatment and prevention of immune diseases.
XX
PS Claim 10; Page 14; 72pp; English.
XX
XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
CC weight TV-marker polypeptides from the present invention. The present
CC invention describes polypeptides (I) for determining the molecular weight
CC of a copolymer (CP), which has an identified molecular weight and an
CC amino acid composition corresponding to the copolymer. The polypeptides
CC of the invention are used as molecular weight markers for glatiramer
CC acetate related tetrapolymers. The polypeptides may also be used for
CC treating and preventing immune diseases in a mammal. Autoimmune diseases
CC diseases. Such diseases include arthritic conditions, demyelinating
CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
CC disease, chronic immune thrombocytopenia purpura, colitis, contact
CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
CC mediated diseases which can be treated include host-versus-graft disease,
CC graft-versus-host disease, and delayed-type hypersensitivity. The
CC polypeptides of the invention have defined molecular weights and physical
CC properties which are analogous to glatiramer acetate molecules, which
CC makes them ideal for use as molecular weight markers
XX
SQ Sequence 77 AA;

Query Match 55.7%; Score 289; DB 3; Length 77;
Best Local Similarity 67.0%; Pred. No. 4.1e-17;
Matches 73; Conservative 1; Mismatches 3; Indels 32; Gaps 4;

QY 1 AKKYAKKAEKAYAKKAAKAEKKAYAKKAEKAYAKKAEKAYAKKAEKAYAKKAEKAYAKKAE 60
DB 1 AKKYAKK-EKAYAKKA-----EKAKKAEKAYAKKAEKAKKA----- 36

QY 61 YKAEAKKYAKKAEKKEKYAAAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 109
DB 37 -KAEAKKYAKKAEKKEKYAAAEK-----YKAEAKKAEKAEKAEKAEKAEKAEKAE 77

RESULT 4
AAY82574
ID AAY82574 standard; peptide; 66 AA.
XX
AC AAY82574;
XX
XX 28-JUL-2000 (first entry)
XX
XX Copolymer molecular weight TV-marker amino acid sequence SEQ ID NO:4.
XX
XX Copolymer; molecular weight marker; TV-marker; immune disease;
KW glatiramer acetate; autoimmune disease; antiarthritic; neuroprotective;
KW osteopathic; immunosuppressive; antithyroid; antiinflammatory;
KW antidiabetic; thyromimetic; haemostatic; antipsoriatic; dermatological;
KW antihaemic; immunosuppressive; demyelinating disease; osteoarthritis;
KW inflammatory condition; multiple sclerosis; rheumatoid arthritis;
KW Crohn's disease; chronic immune thrombocytopenia purpura; colitis;
KW Hashimoto's disease; Graves disease; Guillain-Barre's syndrome; psoriasis;
KW pemphigus vulgaris; systemic lupus erythematosus.
XX
OS Unidentified.
XX
PN WO200018794-A1.

XX PD 06-APR-2000.
XX PF 24-SEP-1999; 99WO-US022402.
XX PR 25-SEP-1998; 98US-0101693P.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (TEVA-) TEVA PHARM USA INC.
XX
XX Gad A, Lis D;
XX
XX WPI; 2000-317499/27.
XX
XX Copolymer 1 related polypeptides used as molecular weight markers for
PT glatiramer acetate and for treatment and prevention of immune diseases.
XX
XX Claim 10; Page 14; 72pp; English.
XX
XX AAY82571 to AAY82577 represent specifically claimed copolymer molecular
CC weight TV-marker polypeptides from the present invention. The present
CC invention describes polypeptides (I) for determining the molecular weight
CC of a copolymer (CP), which has an identified molecular weight and an
CC amino acid composition corresponding to the copolymer. The polypeptides
CC of the invention are used as molecular weight markers for glatiramer
CC acetate related tetrapolymers. The polypeptides may also be used for
CC treating and preventing immune diseases in a mammal. Autoimmune diseases
CC diseases. Such diseases include either cell-mediated or antibody-mediated
CC diseases and inflammatory conditions, e.g. multiple sclerosis, rheumatoid
CC arthritis, osteoarthritis, autoimmune haemolytic anaemia, autoimmune
CC oophoritis, autoimmune thyroiditis, autoimmune uveoretinitis, Crohn's
CC disease, chronic immune thrombocytopenia purpura, colitis, contact
CC syndrome, Hashimoto's disease, idiopathic myxoedema, myasthenia gravis,
CC psoriasis, pemphigus vulgaris, or systemic lupus erythematosus. Mediated-
CC mediated diseases which can be treated include host-versus-graft disease,
CC graft-versus-host disease, and delayed-type hypersensitivity. The
CC polypeptides of the invention have defined molecular weights and physical
CC properties which are analogous to glatiramer acetate molecules, which
CC makes them ideal for use as molecular weight markers
XX
SQ Sequence 66 AA;

Query Match 44.0%; Score 228.5; DB 3; Length 66;
Best Local Similarity 56.9%; Pred. No. 3.8e-12;
Matches 62; Conservative 0; Mismatches 4; Indels 43; Gaps 4;

QY 1 AKKYAKKAEKAYAKKAAKAEKKAYAKKAEKAYAKKAEKAYAKKAEKAYAKKAEKAEKAE 60
DB 1 AKKYAKK-EKAYAKKA-----EKAKKAEKAYAKKAEKAKKA----- 25

QY 61 YKAEAKKYAKKAEKKEKYAAAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 109
DB 26 -KAEAKKYAKKAEKKEKYAAAEK-----YKAEAKKAEKAEKAEKAEKAEKAEKAE 66

RESULT 5
AAR06445
ID AAR06445 standard; protein; 154 AA.
XX
AC AAR06445;
XX
XX 25-MAR-2003 (revised)
XX DT 03-JAN-1991 (first entry)
XX
XX Recombinant copolymer 1-77, myelin basic protein analogue.
DE Recombinant copolymer 1; COP-1-77; myelin basic protein; MBP;
XX immunological activity; autoimmune encephalomyelitis; multiple sclerosis.
XX
XX Synthetic.
XX
XX
```

```

PN EP383620-A.
XX
XX
PD 22-AUG-1990.
XX
XX
PF 16-FEB-1990; 90BP-00301700.
XX
XX
PR 17-FEB-1989; 89US-00312541.
PR 07-FEB-1990; 90US-00473845.
XX
XX
PA (REPK ) REPLIGEN CORP.
XX
XX
PI Cook KS;
XX
XX
DR WPI; 1990-255848/34.
DR N-PSDB; AAQ05664.
XX
XX
PT Producing genes encoding random polymers of aminoacid(s) - for producing
PT recombinant polypeptide(s) with biological and/or immunological activity.
XX
XX
PS Disclosure; Fig 11; 25pp; English.
XX
XX
CC To improve the expression of rCOP-1 polypeptides in E. coli, genes coding
CC for rCOP-1-77 were subcloned from pREV 2.1 to pG3-2deltaN (deposit: 20-
CC NOV-1984 US4691009, NHRB B-15910), a plasmid used to express Protein A.
CC The resulting plasmids encode fusion proteins consisting of beta-
CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue
CC occurs between the Protein A and rCOP-1 sequences, originating from the
CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved
CC from the fusion protein. rCOP-1-77 contains oligonucleotide duplexes
CC encoding the following segments: YKK, EAE, KAK, AAK, and AAA. The N-
CC terminal alanine residue is left behind following CNBr cleavage of the
CC fusion protein. The product prevents or arrests experimental autoimmune
CC encephalomyelitis. They are used to prevent, arrest or control a
CC demyelinating disorder, e.g. multiple sclerosis. They may also be used as
CC additives to hair care products to confer beneficial effects on damaged
CC hair or as supplements for diets deficient in certain amino acids. See
CC also AAQ05665. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX
SQ Sequence 154 AA;

Query Match 40.3%; Score 209; DB 2; Length 154;
Best Local Similarity 53.7%; Pred. No. 3.9e-10;
Matches 72; Conservative 8; Mismatches 20; Indels 34; Gaps 9;

QY 2 KKYAKAEKA-----YAKKAAAEKKKAYA---KKEAKAYKAEAE-----KKKAK-A 44
DB 24 KKKAKAEAEKAKKAKYKKYKKEAEAEKAAKAAKAAKAAAYKKAEAE--AAAEKAKYKKKAKAE 81

QY 45 EAKKYAKEAAKA--KKEAYKAEAKYKAAKAEK-----KEYAAAEAKKAEAA--K 91
DB 82 EYKKKAKAAAEAEYKKEAEAEYKKYKKKAKKAKYKKKAEKAEKAKAAAEAEKAEYK 141

QY 92 AYKAEAAKAAAEKAE 105
DB 142 KYKKEAEK--AKEA 153

RESULT 6
ABO67048
ID ABO67048 standard; protein; 469 AA.
XX
XX
AC ABO67048;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polypeptide seqid 13565.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
XX
XX
OS Klebsiella pneumoniae.
XX
XX
PN US6610836-B1.

XX
XX
26-AUG-2003.
XX
XX
27-JAN-2000; 2000US-00489039.
XX
XX
29-JAN-1999; 99US-0117747P.
XX
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
XX
PI Breton GL, Osborne M;
XX
XX
DR WPI; 2003-895346/82.
DR N-PSDB; ABD00619.
XX
XX
PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
PT preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX
PS Disclosure; SEQ ID NO 13565; 932pp; English.
XX
XX
CC The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention
XX
XX
SQ Sequence 469 AA;

Query Match 37.8%; Score 196; DB 7; Length 469;
Best Local Similarity 51.2%; Pred. No. 1.5e-08;
Matches 62; Conservative 12; Mismatches 35; Indels 12; Gaps 4;

QY 1 AKKYAKAEKAYKAKAKAEK---KAYAKKEAKAYKAAEAKKAKAEAK----KYAKEA 53
DB 153 AKEQKQAEAEAAKAAKAAKAKADAKAEQAEAAKAAAEAKAKADAKAQAKAAEQAAAKAA 212

QY 54 AKAKKEAYKAEAKYKAKA---AKAEKKEVYAAAEAKKAEAAKAYKA--EAKAKAAKEAAVE 108
DB 213 ADAKQAEAEAAKAAAEAKKQAEAEAAKAAAEAKKAEAAKAAAEAKKAEAAKQAEAEAKQ 272

QY 109 A 109
DB 273 A 273

RESULT 7
ABU40185
ID ABU40185 standard; protein; 372 AA.
XX
XX
AC ABU40185;
XX
XX
DT 19-JUN-2003 (first entry)
XX
XX
DE Protein encoded by Prokaryotic essential gene #25712.
XX
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX
OS Pseudomonas putida.
XX
XX
PN WO200277183-A2.
XX
XX
PD 03-OCT-2002.
XX
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX
PN (ELIT-) ELITRA PHARM INC.

```


QY 101 AAKEAA 106
 Db 285 AAAXAA 290

RESULT 9
 ABU28559
 ID ABU28559 standard; protein; 421 AA.
 XX AC ABU28559;
 XX 19-JUN-2003 (first entry)
 XX Protein encoded by Prokaryotic essential gene #14086.
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Escherichia coli.
 OS
 XX WO200277183-A2.
 XX PD 03-OCT-2002.
 XX PF 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA32429.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids, required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 56483; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC on a gene on which the test compound that inhibits proliferation of an
 CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *k. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 421 AA;
 Query Match 36.5%; Score 189.5; DB 6; Length 421;
 Best Local Similarity 53.0%; Pred. No. 4.7e-08;
 Matches 61; Conservative 11; Mismatches 32; Indels 11; Gaps 5;
 QY 1 AKYAKKAB-----KAYAKKAAKAEKKYAKKAYKAAKAAKAAKAAKAAKAAK 55
 Db 120 ABEAAKQAEKQKQAEAAKAAADAKAAKAEADAKA--AEAAKKAADAKKAAEAAK 177
 QY 56 AKKEAY-KAEAKYKAAKAAKAEKVAEAAKAK---AEAAKAYKAAKAAKAAKAA 106
 Db 178 AAARQKAAEAAAALKKKAAEAAEAAAEAAEAAEAAEAAEAAEAAEAAEAAEAA 232

RESULT 10
 AAR06446
 ID AAR06446 standard; protein; 106 AA.
 XX AC AAR06446;
 XX 25-MAR-2003 (revised)
 DT 03-JAN-1991 (first entry)
 XX Recombinant copolymer 1-19, myelin basic protein analogue.
 DE Recombinant copolymer 1; COP-1-19; myelin basic protein; MBP;
 KW immunological activity; autoimmune encephalomyelitis; multiple sclerosis.
 XX Synthetic.
 OS EP383620-A.
 PN 22-AUG-1990.
 XX 16-FEB-1990; 90EP-00301700.
 XX 17-FEB-1989; 89US-00312541.
 PR 07-FEB-1990; 90US-00473845.
 XX (REPK) REPLIGEN CORP.
 PA Cook KS;
 PI WPI; 1990-255848/34.
 DR N-PSDB; AAQ06446.
 XX Producing genes encoding random polymers of aminoacid(s) - for producing
 PT recombinant polypeptide(s) with biological and/or immunological activity.
 XX Disclosure; Fig 12; 25pp; English.
 XX To improve the expression of rCOP-1 polypeptides in *E. coli*, genes coding
 CC for rCOP-1-19 were subcloned from pREV 2.1 to pBG3-2deltaN (deposit: 20-
 CC NOV-1984 US4691009, NRRL B-15910), a plasmid used to express Protein A.
 CC The resulting plasmids encode fusion proteins consisting of beta-
 CC glucuronidase, Protein A, and rCOP-1 sequences. A methionine residue
 CC occurs between the Protein A and rCOP-1 sequences, originating from the
 CC 5' linker sequence, in order that the COP-1 polypeptide may be cleaved
 CC from the fusion protein. rCOP-1-19 contains oligonucleotide duplexes
 CC including the following segments: YKK, AAE, EKA, EKA, YEA, AKA KEA,
 CC and AAA. The N-terminal alanine residue is left behind following CNBr
 CC cleavage of the fusion protein. The product prevents or arrests
 CC experimental autoimmune encephalomyelitis. They are used to prevent,
 CC arrest or control a demyelinating disorder, e.g. multiple sclerosis. They
 CC may also be used as additives to hair care products to confer beneficial
 CC effects on damaged hair or as supplements for diets deficient in certain
 CC amino acids. See also AAQ05664. (Updated on 25-MAR-2003 to correct PA
 CC field.)

This Page Blank (uspto)

Run on: December 14, 2004, 05:52:22 ; Search time 270.43 Seconds
(without alignments)
143.965 Million cell updates/

Sequence: 1 AKKYAKKAEKAYAKKAKAAK.....AKAYKAEAAKAAKEAAAYEA 109

Scoring table: $\text{E} = 0.00002$
Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 15855576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
2: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgm2_6/ptodata/1/pubpaa/US05_NEW_PUB.pcp.*
4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp.*
5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp.*
6: /cgm2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp.*
7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp.*
8: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
9: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp.*
10: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp.*
11: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp.*
12: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp.*
13: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
14: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
15: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp.*
16: /cgm2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp.*
17: /cgm2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp.*
18: /cgm2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp.*
19: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp.*
20: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			Description	
	Score	Match	Length	ID	Description
1	519	100.0	109	9	US-09-816-989A-7
2	312.5	60.2	86	9	US-09-816-989A-6
3	289	55.7	77	9	US-09-816-989A-5
4	228.5	44.0	66	9	US-09-816-989A-4
5	190.5	36.7	372	15	US-10-282-122A-68109
6	190	36.6	428	15	US-10-282-122A-55748
7	189.5	36.5	421	15	US-10-282-122A-56483
8	184.5	35.5	407	15	US-10-282-122A-75047
9	181.5	35.0	323	15	US-10-282-122A-59321
10	180.5	34.8	56	9	US-09-816-989A-3
11	179	34.5	347	14	US-10-127-032-120
12	179	34.5	347	15	US-10-282-122A-66237
13	178	34.3	223	13	US-10-051-643-201

14	178	34.3	223	14	US-10-205-979-52	Sequence 52, Appl
15	178	34.3	376	15	US-10-282-122A-75772	Sequence 75772, A
16	178	34.3	388	15	US-10-282-122A-78190	Sequence 78190, A
17	174	33.4	387	15	US-10-282-122A-72645	Sequence 72645, A
18	173.5	33.4	389	15	US-10-282-122A-67145	Sequence 67145, A
19	163.5	31.4	438	15	US-10-282-122A-452664	Sequence 45264, A
20	163	31.4	336	15	US-10-282-122A-69962	Sequence 69962, A
21	163	31.4	376	14	US-10-156-761-9889	Sequence 9889, Ap
22	162	31.2	356	10	US-09-820-843A-27	Sequence 27, App
23	162	31.2	356	15	US-10-282-122A-77342	Sequence 77342, A
24	160.5	30.9	528	14	US-10-229-567-40	Sequence 40, Appl
25	160.5	30.9	226	14	US-10-229-567-32	Sequence 32, Appl
26	160.5	30.9	226	17	US-10-733-878-381	Sequence 381, App
27	159.5	30.7	234	14	US-10-262-209-1	Sequence 1, Appli
28	159.5	30.7	234	16	US-10-240-430-2	Sequence 2, Appli
29	158	30.4	218	14	US-10-229-567-4	Sequence 4, Appli
30	158	30.4	219	17	US-10-733-878-489	Sequence 489, App
31	156.5	30.2	298	15	US-10-425-114-56061	Sequence 56061, A
32	156.5	30.2	303	17	US-10-425-115-297103	Sequence 297103, A
33	155.5	30.0	507	16	US-10-726-692-50	Sequence 50, Appl
34	155.5	30.0	717	16	US-10-726-692-36	Sequence 36, Appl
35	154.5	29.8	130	14	US-10-262-209-2	Sequence 2, Appli
36	154.5	29.8	130	16	US-10-240-430-5	Sequence 5, Appli
37	154.5	29.8	309	16	US-10-437-963-186290	Sequence 186390, A
38	154.5	29.8	309	10	US-09-820-843A-24	Sequence 24, Appl
39	153.5	29.6	369	10	US-09-820-843A-95	Sequence 95, Appl
40	153.5	29.6	369	15	US-10-282-122A-76514	Sequence 76514, A
41	153.5	29.6	372	10	US-09-820-843A-8	Sequence 8, Appli
42	153.5	29.6	372	16	US-10-467-421-16	Sequence 16, Appl
43	153.5	29.6	926	16	US-10-437-963-193381	Sequence 193381, A
44	153	29.4	329	15	US-10-282-122A-67699	Sequence 67699, A
45	152.5	29.4	222	14	US-10-229-567-3	Sequence 3, Appli

ALIGNMENTS

```

RESULT 1
US-09-816-989A-7
; Sequence 7, Application US/09816989A
; Patent No. US20020115103A1
; GENERAL INFORMATION:
; APPLICANT: Gad, Alexander
; APPLICANT: Lis, Doris
; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARKERS
; TITLE OF INVENTION: AND FOR THERAPEUTIC USE
; FILE REFERENCE: 2609/60807-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/816,989A
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/101,693
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/22402
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide
US-09-816-989A-7

```

Query Match	100.0%;	Score 519;	DB 9;	Length 109;
Best Local Similarity	100.0%;	Pred. No. 1.1e-33;		
Matches 109; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AKKYAKKAKYAKKAKAAKEKKYAKKEAKYKAAEAKKKAKAAEKYKAAKAAKKEA	60	
Db	1	AKKYAKKAKYAKKAKAAKEKKYAKKEAKYKAAEAKKKAKAAEKYKAAKAAKKEA	60	
Qy	61	YKAAEKKYAKAAKEKKEYYAAAEAKKAAEAKYKAAEAAKAAKAAKAAEAA	109	

Db 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

RESULT 2

US-09-816-989A-6

; Sequence 6, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; TITLE OF INVENTION: AND FOR THERAPEUTIC USE

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 6

; LENGTH: 86

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-6

Query Match 60.2%; Score 312.5; DB 9; Length 86;

Best Local Similarity 72.1%; Pred. No. 1.3e-17;

Matches 80; Conservative 1; Mismatches 3; Indels 27; Gaps 5;

QY 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKAEKAYKAEAAEAKKKA----- 36

QY 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 37 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 86

RESULT 3

US-09-816-989A-5

; Sequence 5, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; TITLE OF INVENTION: AND FOR THERAPEUTIC USE

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 5

; LENGTH: 77

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-5

Query Match 55.7%; Score 289; DB 9; Length 77;

Best Local Similarity 67.0%; Pred. No. 7.9e-16;

Matches 73; Conservative 1; Mismatches 3; Indels 32; Gaps 4;

QY 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKAEKAYKAEAAEAKKKA----- 36

QY 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 37 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 77

RESULT 4

US-09-816-989A-4

; Sequence 4, Application US/09816989A

; Patent No. US20020115103A1

; GENERAL INFORMATION:

; APPLICANT: Gad, Alexander

; APPLICANT: Lis, Doris

; TITLE OF INVENTION: COPOLYMER 1 RELATED POLYPEPTIDES FOR USE AS MOLECULAR WEIGHT MARK

; TITLE OF INVENTION: AND FOR THERAPEUTIC USE

; FILE REFERENCE: 2609/60807-A-PCT-US

; CURRENT APPLICATION NUMBER: US/09/816,989A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: 60/101,693

; PRIOR FILING DATE: 1998-09-25

; PRIOR APPLICATION NUMBER: PCT/US99/22402

; PRIOR FILING DATE: 1999-09-24

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 4

; LENGTH: 66

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic Peptide

US-09-816-989A-4

Query Match 44.0%; Score 228.5; DB 9; Length 66;

Best Local Similarity 56.9%; Pred. No. 3.6e-11;

Matches 62; Conservative 0; Mismatches 4; Indels 43; Gaps 4;

QY 1 AKKYAKKAEKAYAKKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 60

Db 1 AKKYAKK-EKAYAKKA-----EKAACKA----- 25

QY 61 YKAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 109

Db 26 -KAEAKKYAKAAKAEKKEYAAAEAKKAEAAKAYKAEAAKAAAEAAAYEA 66

RESULT 5

US-10-282-122A-68109

; Sequence 68109, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match          36.7%; Score 190.5; DB 15; Length 372;
Best Local Similarity 50.4%; Pred. No. 2e-07;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

QY 5 AKKAEKAYAKKAAK---EKKAYAKKAEKAYKAAE-----AKKAEKAEKAYKAAE 54
Db 118 AKKAEDA-AKAAEAAKAAKAAEAKKADEAKKAAEKQOADIJAKKAEDEAKKAAEBA 176

QY 55 K-AKKEAYKAEKAYKAAKAEKKEEYAAAEAKKAEAAKAYKAAEAKAAEAAVEA 109
Db 177 KKAABEAAKKAEDAKKAAEBAKKAEDAKKAAEDAKKAAEAAEDAKKAAEAAKKAADA 233

RESULT 6
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match          36.7%; Score 190.5; DB 15; Length 372;
Best Local Similarity 50.4%; Pred. No. 2e-07;
Matches 59; Conservative 14; Mismatches 31; Indels 13; Gaps 4;

QY 5 AKKAEKAYAKKAAK---EKKAYAKKAEKAYKAAE-----AKKAEKAEKAYKAAE 54
Db 118 AKKAEDA-AKAAEAAKAAKAAEAKKADEAKKAAEKQOADIJAKKAEDEAKKAAEBA 176

QY 55 K-AKKEAYKAEKAYKAAKAEKKEEYAAAEAKKAEAAKAYKAAEAKAAEAAVEA 109
Db 177 KKAABEAAKKAEDAKKAAEBAKKAEDAKKAAEDAKKAAEAAEDAKKAAEAAKKAADA 233

RESULT 6
US-10-282-122A-55748
; Sequence 55748, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 68109
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-68109

Query Match          36.6%; Score 190; DB 15; Length 428;
Best Local Similarity 50.8%; Pred. No. 2.5e-07;
Matches 64; Conservative 8; Mismatches 32; Indels 22; Gaps 5;

QY 1 AKKYA-----KKAEKAYAKKAAKAEKAYKAAEAKKAEKAYKAAEAKKAEKAAE 47
Db 167 AKKAADAQKKAEEAAKKAADAQKKAEEAAKKA--AADQKKAEEAAKKAEEAEK 224

QY 48 KYAEAAKAEKAYKAAEAKKAEKAYKAAEAKKAEKAYKAAEAKKAEKAAEAKKAE 100
Db 225 KAAEAAKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAKKAAEAK 284

QY 101 AAKEAA 106
Db 285 AAKKAA 290

RESULT 7
US-10-282-122A-56483
; Sequence 56483, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

	Matches	66;	Conservative	11;	Mismatches	25;	Indels	17;	Gaps	8
	; NUMBER OF SEQ ID NOS: 78614 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 56483 ; LENGTH: 421 ; TYPE: PRT ; ORGANISM: Escherichia coli US-10-282-122A-56483									
	Query Match	36.5%;	Score	189.5;	DB 15;	Length	421;			
	Best Local Similarity	53.0%;	Pred.	No. 2.7e-07;						
	Matches	61;	Conservative	11;	Mismatches	33;	Indels	11;	Gaps	5;
OY	1 AKKYAKKAE-----KAYAKKAKEKKYAKKAYKAAEAKKAYKAAEKYAKKAAEAKKAYAK-AAK 55	: : : :	:	:	:	:	:	:	:	:
Dd	120 ABEAAQQAELKQAEAEAAKADAKAKAEADA--ABEAACKAAADAKKKAAEA 177	: : : :	:	:	:	:	:	:	:	:
OY	56 AKKEAY-KAEAKKYAKKAKEKKYAAAENK---AEAAYKAAEAKAAEAA 106	: : : :	:	:	:	:	:	:	:	:
Dd	178 AAEEAQKAAEAAAAALKKKAEEEEAAAAEARKGAATEAAERKAAEAKKAAEAKAAEKA 232	: : : :	:	:	:	:	:	:	:	:
	RESULT 8 US-10-282-122A-75047 ; Sequence 75047, Application US/10282122A ; Publication No. US20040029129A1 ; GENERAL INFORMATION: ; APPLICANT: Wang, Liangsu ; APPLICANT: Zamudio, Carlos ; APPLICANT: Malone, Cheryl ; APPLICANT: Haselbeck, Robert ; APPLICANT: Ohlsen, Kari ; APPLICANT: Zyskind, Judith ; APPLICANT: Wall, Daniel ; APPLICANT: Trawick, John ; APPLICANT: Carr, Grant ; APPLICANT: Yamamoto, Robert ; APPLICANT: Forsyth, R. ; APPLICANT: Xu, H. ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms ; FILE REFERENCE: ELITRA.034A ; CURRENT FILING DATE: 2003-02-20 ; PRIOR APPLICATION NUMBER: 60/191,078 ; PRIOR FILING DATE: 2000-03-21 ; PRIOR APPLICATION NUMBER: 60/206,848 ; PRIOR FILING DATE: 2000-05-23 ; PRIOR APPLICATION NUMBER: 60/207,727 ; PRIOR FILING DATE: 2000-05-26 ; PRIOR APPLICATION NUMBER: 60/230,335 ; PRIOR FILING DATE: 2000-09-06 ; PRIOR APPLICATION NUMBER: 60/230,347 ; PRIOR FILING DATE: 2000-09-09 ; PRIOR APPLICATION NUMBER: 60/242,578 ; PRIOR FILING DATE: 2000-10-23 ; PRIOR APPLICATION NUMBER: 60/253,625 ; PRIOR FILING DATE: 2000-11-27 ; PRIOR APPLICATION NUMBER: 60/257,931 ; PRIOR FILING DATE: 2000-12-22 ; PRIOR APPLICATION NUMBER: 60/267,636 ; PRIOR FILING DATE: 2001-02-09 ; PRIOR APPLICATION NUMBER: 60/269,308 ; PRIOR FILING DATE: 2001-02-16 ; Remaining Prior Application data removed - See File Wrapper or PALM. ; NUMBER OF SEQ ID NOS: 78614 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 75047 ; LENGTH: 407 ; TYPE: PRT ; ORGANISM: Salmonella typhimurium US-10-282-122A-75047									
	Query Match	35.5%;	Score	184.5;	DB 15;	Length	407;			
	Best Local Similarity	55.5%;	Pred.	No. 6.4e-07;						

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 05:47:53 ; Search time 30.5844 Seconds
(without alignments)
236.351 Million cell updates/sec

Title: US-10-792-311-7
Perfect score: 519
Sequence: 1 AKKYAKKAEKAYAKKAAK.....AKYKAEAKAAKEAAEYEA 109

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6C COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	519	100.0	109	4 US-09-405-743A-7	Sequence 7, Appli
2	312.5	60.2	86	4 US-09-405-743A-6	Sequence 5, Appli
3	289	55.7	77	4 US-09-405-743A-5	Sequence 4, Appli
4	228.5	44.0	66	4 US-09-405-743A-4	Sequence 13565, A
5	196	37.8	469	4 US-09-489-039A-13565	Sequence 3, Appli
6	180.5	34.8	56	4 US-09-405-743A-3	Sequence 64, Appl
7	179.5	34.6	100	2 US-08-460-890A-64	Sequence 64, Appl
8	179.5	34.6	100	3 US-08-167-641C-64	Sequence 64, Appl
9	179.5	34.6	100	3 US-08-460-971A-64	Sequence 64, Appl
10	179.5	34.6	100	3 US-08-462-040-64	Sequence 29581, A
11	179	34.5	407	4 US-09-252-991A-29581	Sequence 201, App
12	178	34.3	223	3 US-09-095-855-201	Sequence 5390, Ap
13	178	34.3	223	4 US-09-205-426-201	Sequence 39148, A
14	169	32.6	361	4 US-09-543-681A-5390	Sequence 54365, A
15	164	31.6	171	4 US-09-270-767-39148	Sequence 8, Appli
16	164	31.6	171	4 US-09-270-767-54365	Sequence 8, Appli
17	162.5	31.3	643	2 US-08-216-894-8	Sequence 40, Appl
18	162.5	31.3	643	3 US-09-115-746-8	Sequence 40, Appl
19	160.5	30.9	158	3 US-09-041-889-40	Sequence 32, Appl
20	160.5	30.9	158	4 US-09-417-264-40	Sequence 32, Appl
21	160.5	30.9	226	3 US-09-041-889-32	Sequence 6321, Ap
22	160.5	30.9	226	4 US-09-417-264-32	Sequence 5169, Ap
23	160.5	30.9	468	4 US-09-328-352-6321	Sequence 10, Appl
24	160.5	30.8	214	4 US-09-328-352-5169	Sequence 10, Appl
25	158.5	30.5	472	2 US-08-216-894-10	Sequence 2, Appli
26	158.5	30.5	472	3 US-09-115-746-10	
27	158.5	30.5	564	2 US-08-216-894-2	

28	158.5	30.5	564	3 US-09-115-746-2	Sequence 2, Appli
29	158	30.4	218	3 US-09-041-889-4	Sequence 4, Appli
30	158	30.4	218	3 US-08-837-058-4	Sequence 4, Appli
31	158	30.4	218	4 US-09-417-264-4	Sequence 3, Appli
32	152.5	29.4	222	3 US-09-041-889-3	Sequence 3, Appli
33	152.5	29.4	222	3 US-08-837-058-3	Sequence 3, Appli
34	152.5	29.4	222	4 US-09-417-264-3	Sequence 3, Appli
35	151	28.1	316	4 US-09-252-991A-32957	Sequence 32957, A
36	150	28.5	1507	3 US-08-929-329-5	Sequence 5, Appli
37	148	28.5	207	4 US-09-489-039A-13743	Sequence 13743, A
38	144	27.7	399	4 US-09-252-991A-22853	Sequence 22853, A
39	144	27.7	434	1 US-08-097-830E-3	Sequence 3, Appli
40	144	27.7	434	2 US-08-456-112B-3	Sequence 3, Appli
41	143.5	27.6	921	4 US-09-543-681A-5734	Sequence 5734, Ap
42	143	27.6	214	3 US-09-041-889-27	Sequence 27, Appl
43	143	27.6	214	4 US-09-417-264-27	Sequence 27, Appl
44	141	27.2	100	2 US-08-460-890A-62	Sequence 62, Appl
45	141	27.2	100	3 US-08-167-641C-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-405-743A-7
; Sequence 7, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-7

Query Match 100.0%; Score 519; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 7.1e-37;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AKKYAKKAEKAYAKKAAKKEKKAYAKKEKKAYAKKAAEAKKKAKAEAKKAKKAAKKEA 60
Db 1 AKKYAKKAEKAYAKKAAKKEKKAYAKKEKKAYAKKAAEAKKKAKAEAKKAKKAAKKEA 60
QY 61 YKAEAKKAKAKAEKKEVAAAEAKKAEAKKAYAKKAAKAAKAAKAAKAAKAAKAAEYEA 109
Db 61 YKAEAKKAKAKAEKKEVAAAEAKKAEAKKAYAKKAAKAAKAAKAAKAAKAAKAAEYEA 109

RESULT 2
US-09-405-743A-6
; Sequence 6, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: PEPTIDE
; US-09-405-743A-6
SYNTHETIC

```

```

Query Match      60.2%; Score 312.5; DB 4; Length 86;
Best Local Similarity 72.1%; Pred. No. 1e-19;
Matches 80; Conservative 1; Mismatches 3; Indels 27; Gaps 5;

Qy 1 AKGYAKAEAYAKAKAAKAEKKAYAKKEAKAYKAAEAKKAAEAKKYAKAEAKAKKEA 60
      |||||
Db 1 AKGYAKK- EKAYAKKA-----EKAKEAKAEAKAYKAAEAKKKA-----36

```

```

RESULT 3
US-09-405-743A-5
; Sequence 5, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-5
SYNTHETIC

```

```
Query Match      55.7%; Score 289; DB 4; Length 77;  
Best Local Similarity 67.0%; Pred. No. 8.2e-18;  
Matches          73; Conservative 1; Mismatches 3; Indels 32; Gaps 4;
```

```
Qy   1 AKTYAKAEKAYAKAKAAEKKYAYKEAKYTAAEAKKYKAEEAKKYAKEAAKKEA 60  
      ||||| ||||| ||||| :||| ||||| ||||| ||||| |||||  
Db   1 AKTYAKK-EKAYAKKA-----EKAAKAEAKYKAAEAACKKA----- 36
```

```
Qy   61 YKAEAKYKAAKAEKKEYYAAAEEKYAEAAKAYTAEAAKAAKAAKAAVEA 109
```

```
Db   37 -KAEAKYKAAKAEKKEYYAAAEAK-----YKAEAAKAAKAAKAAVEA 77
```

```

RESULT 4
US-09-405-743A-4
; Sequence 4, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: PEPTIDE
US-09-405-743A-4
SYNTHETIC

```

```

Query Match      44.0%; Score 228.5; DB 4; Length 66;
Best Local Similarity 56.9%; Pred. No. 7.9e-13;
Matches 62; Conservative 0; Mismatches 4; Indels 43; Gaps 4;

Qy      1  AKYAKAEKAYAKKAKAAKEKKYAYAKEBAKAYKAAEAKKKAKAEAKKYAKAEAAKAKEA 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      1  AKTYAKK-EKAYAKAKKA-----EAKAKKA----- 25

Qy      61  KYAEAKKYAKAAAEKKEYAAAAEKKAEBAKAYKAEAAKAAKAAKAEAA 109
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      26  -KAEAKKYAKAAAEKKEYAAAAEAK-----KYAEAAKAAKAAKAEAA 66
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 5
US-09-489-039A-13565
; Sequence 13565, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13565
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13565

Query Match	37.8%;	Score 196;	DB 4;	Length 469;
Best Local Similarity	51.2%;	Pred. No. 2.9e-09;		
Matches	62;	Conservative 12;	Mismatches 35;	Indels 12; Gaps 4
Qy	1	AKKVAAKEAYAKKAAAKEK---	KAYAKKEAYKAAAEAKKKAAEAK---	KYAKEA 53
Dd	153	AKFQQKQAEAAKAAAATAKADAQ	KEAQEEAAAKAAAEAKAKADAKQAQKAQEAQAAKAA	212
Qy	54	AKAKCEAYKAEAKKYAKA---	AKAEKCEYAAAAEKAAEAAKAYKA---	EAAKAAAEKAAAYE 108
Dd	213	ADAKKQAEAAAAKAAAEAKKQAE	AEEAAKAAAEAAQCKAAEAAAKKAAQQAQQAQQAQKQ	272
Qy	109	A 109		
Dd	273	A 273		

```

RESULT 6
US-09-405-743A-3
; Sequence 3, Application US/09405743A
; Patent No. 6514938
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co., Ltd.
; TITLE OF INVENTION: GLATIRAMER ACETATE MOLECULAR WEIGHT MARKERS
; FILE REFERENCE: 60807-A
; CURRENT APPLICATION NUMBER: US/09/405,743A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; - OTHER INFORMATION: PEPTIDE
US-09-405-743A-3
SYNTHETIC

```

Query Match	34.8%	Score 180.5;	DB 4;	Length 56;
Best Local Similarity	45.9%	Pred. No. 6.9e-09;		


```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29581
; LENGTH: 407
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29581

Query Match 34.5%; Score 179; DB 4; Length 407;
Best Local Similarity 49.1%; Pred. No. 6.7e-08;
Matches 54; Conservative 17; Mismatches 29; Indels 10; Gaps 4;

Qy 6 KKAEEKAYAKAKAAKEK-----KAYAKKEAKYKAAEAKKKAKAEAKKAKAEAKKAKE 59
Db 159 QKLEQQQVAAAKAAEQKKADEARKAEAKQKAAEAACKADEAKKAAEAKEAKKAAEQKQADIATKCR 218
Qy 60 AYKAEAKYKAKAAKAEKKEKAAEAACKAAEAAYKAAEAAYKAAEAAYEA 109
Db 219 A-EDEAKK--KAAEDAKK--AAEDAKKAAEAEEKKAAEAANKKAAEA 264

RESULT 12
US-09-095-855-201
; Sequence 201, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565

```

```
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5390
; LENGTH: 361
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5390

Query Match      32.6%; Score 169; DB 4; Length 361;
Best Local Similarity 44.0%; Pred. No. 4e-07;
Matches 51; Conservative 16; Mismatches 37; Indels 12; Gaps 2;

QY 1 AKKYAKAEKAYAKKAAKAEKAYAKKAEKAYAKA-----AAKAKKAEAKKAYAK 51
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 123 AKEAAEQRRREATAAAKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 182
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 52 EAAKAKAEKAYAKKAEKAYAKA---AKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 104
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 183 LAAKQAEAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEK 238

RESULT 15
US-09-270-767-39148
; Sequence 39148, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39148
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-39148

Query Match      31.6%; Score 164; DB 4; Length 171;
Best Local Similarity 42.9%; Pred. No. 5e-07;
Matches 57; Conservative 20; Mismatches 30; Indels 26; Gaps 6;

QY 2 KKYAKAEKAYAKKAAKAEKK-----AYAKKE--AKAYKAEAKKK-----AKAEAK 47
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 22 KKAQKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 81
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 48 KYAKAA-----KAKKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 96
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 82 KAAEEALKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 140
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
QY 97 AAKAAAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 109
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 141 AAQKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAEKAE 153
```

Search completed: December 14, 2004, 05:50:19
Job time : 31.7272 secs